



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 164310

TO: Jon E Angell
Location: 2d20 / 2c18
Friday, September 02, 2005
Art Unit: 1635
Phone: 571-272-0756
Serial Number: 09 / 972916

From: Jan Delaval
Location: Biotech-Chem Library
Remsen 1a51
Phone: 571-272-2504

jan.delaval@uspto.gov

Search Notes

This Page Blank (uspto)

From: Chan, Christina
Sent: Tuesday, August 30, 2005 11:10 AM
To: Angell, Jon E; STIC-Biotech/ChemLib
Subject: RE: RUSH Sequence Search Request 09/972,916

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Angell, Jon E
Sent: Tuesday, August 30, 2005 11:24 AM
To: Chan, Christina
Subject: RUSH Sequence Search Request 09/972,916

I would like to request a RUSH sequence search because this case is an after final amended case that was forwarded to me 8/29/2005.

If approved, please forward the request to STIC and CC a copy to me, THANKS!

SEARCH REQUEST FORM

Scientific and Technical Information Center

Examiner# : 78697
Art Unit : 1635
Phone Number: 571-272-0756
Date: 8/30/2005
Serial Number: 09/972,916 (THULE et al.)
Mailbox & Bldg/Room Location: REMSEN 2C18
Results Format Preferred (circle): Paper

I would like to have a **standard** search of the nucleic acid databases including the pending applications databases (i.e., an **interference** search) performed using the following SEQ. ID NOs. from application : 09/972,916

SEQ ID NO: 1 (nucleic acid 51 nucleotides in length)
SEQ ID NO: 2 (nucleic acid 219 nucleotides in length)
SEQ ID NO: 3 (nucleic acid 270 nucleotides in length)
SEQ ID NO: 4 (nucleic acid 321 nucleotides in length)
SEQ ID NO: 5 (nucleic acid 372 nucleotides in length)
SEQ ID NO: 6 (nucleic acid 423 nucleotides in length)

Please note that the above sequence are gene expression regulatory sequences and the claims encompass nucleic acid sequences comprising any 1 or more of the above sequences.

STAFF USE ONLY

Searcher: Q
Searcher Phone: 2- 2504
Date Searcher Picked up: 8/30/05
Date Completed: 9/2/05
Searcher Prep/Rev. Time: 15
Online Time: 15

Type of Search

NA#: ✓ AA#:
Interference: ✓ SPDI:
S/L: Oligomer:
Encode/Transl:
Structure#: Text:
Inventor: Litigation:

Vendors and cost where applicable

STN:
DIALOG:
QUESTEL/ORBIS:
LEXIS/NEXIS:
SEQUENCE SYSTEM: ✓
WWW/Internet:
Other(Specify):

This Page Blank (uspto)

Please contact me by phone or email if you have any questions.

Thanks,
Eric

J. Eric Angell
Art Unit 1635
Office: REMSEN 2D20
mailbox: REM 2C18
571-272-0756

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 12:52:03 ; Search time 232.225 Seconds
(without alignment)
5582.626 Million cell updates/sec

Title: US-09-972-916b-2

Perfect score: 219
Sequence: 1 tcacagcgaacaaactta.....ttccgctactagctagccgc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: _geneseqn1980s:*
2: _geneseqn1990s:*
3: _geneseqn2000s:*
4: _geneseqn2001as:*
5: _geneseqn2001bs:*
6: _geneseqn2002as:*
7: _geneseqn2002bs:*
8: _geneseqn2003as:*
9: _geneseqn2003bs:*
10: _geneseqn2003cs:*
11: _geneseqn2003ds:*
12: _geneseqn2004as:*
13: _geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	219	100.0	219	6	ABX15372 Rat insul
2	219	100.0	270	6	ABX15373 Rat insul
3	219	100.0	321	6	ABX15374 Rat insul
4	219	100.0	372	6	ABX15375 Rat insul
5	219	100.0	423	6	ABX15376 Rat insul
6	219	100.0	474	6	ABX15377 Rat insul
7	219	100.0	525	6	ABX15378 Rat insul
8	219	100.0	576	6	ABX15379 Rat insul
9	219	100.0	627	6	ABX15380 Rat insul
10	219	100.0	678	6	ABX15381 Rat insul
11	219	100.0	729	6	ABX15382 Rat insul
12	219	100.0	780	6	ABX15383 Rat insul
13	219	100.0	831	6	ABX15384 Rat insul
14	219	100.0	882	6	ABX15385 Rat insul
15	219	100.0	933	6	ABX15386 Rat insul
16	219	100.0	984	6	ABX15387 Rat insul
17	219	100.0	1035	6	ABX15388 Rat insul
18	219	100.0	1086	6	ABX15389 Rat insul
19	219	100.0	1137	6	ABX15390 Rat insul
20	219	100.0	1188	6	ABX15391 Rat insul

ALIGNMENTS

21	32	14.6	76180	13	ABD33385	ABD33385 Human can
22	31.6	14.4	548	13	ADO51481	ADO51481 Novel can
23	31.6	14.4	927	6	ABL57730	ABL57730 Human bbg
24	31.6	14.4	1125	6	AAF94440	AAF94440 Human hyd
25	31.6	14.4	1129	6	AAD33646	AAD33646 Human TRI
26	31.6	14.4	1288	10	ADC87210	ADC87210 Human GPC
27	31.6	14.4	1344	6	ABL57731	ABL57731 Human bbg
28	31.6	14.4	1348	6	ABX71176	ABX71176 Human hum
29	31.6	14.4	1513	6	AD116493	AD116493 Human NOV
30	31.6	14.4	1550	4	AAF94450	AAF94450 Human hyd
31	31.6	14.4	1803	4	AAE94689	AAE94689 Human ful
32	31.6	14.4	1803	12	ADL31671	ADL31671 Full leng
33	31.6	14.4	117750	13	ABD33653	ABD33653 Human can
34	31.4	14.3	1134	8	ACA45381	ACA45381 Prokaryot
35	31.2	14.2	668	6	ABQ53903	ABQ53903 Oligonuc1
36	31.2	14.2	668	6	ABQ53902	ABQ53902 Oligonuc1
37	31.2	14.2	955	3	AAF10246	AAF10246 Fusedium
38	31.2	14.2	6226	6	ABK31391	ABK31391 Signal tr
39	31.2	14.2	6226	6	ABL70338	ABL70338 Chemical tr
40	31.2	14.2	6226	6	AAE61293	AAE61293 Human gen
41	31.2	14.2	406	2	AAV90337	AAV90337 EST clone
42	31	14.2	2190	10	ABZ84770	ABZ84770 Human cal
43	31	14.2	2190	11	ABD21000	ABD21000 Human cal
44	31	14.2	2190	13	ADN60285	ADN60285 Human cal
45	31	14.2	2247	5	AAE44950	AAE44950 CDNA enco

RESULT 1

ABX15372 standard; DNA; 219 BP.

ABX15372;

17-APR-2003 (first entry)

Rat insulin-sensitive element (ISE) DNA.

Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1; insulin-sensitive element; ISE; basal promoter; hypoglycaemia; insulin; insulin-like growth factor binding protein-1; hypoglycaemia; glucose; fat catabolism; protein catabolism; carbohydrate; glucocorticoid; fat; glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis; hepatocyte; hepatoma; cellular protein degradation; antidiabetic; intracellular hormone receptor; insulin regulator construct; anabolic.

Rattus norvegicus.

US2002107198-A1.

08-AUG-2002.

10-OCT-2001; 2001US-00972916.

11-OCT-2000; 2000US-0239113P.

(THUL/) THUL P M.

Thule PM;

WPI; 2002-674190/72.

New insulin regulator cassette, useful e.g. for treating diabetes, provides specific, glucose-inducible transgenic expression of insulin in liver cells.

Claim 8; Page 13; 37pp; English.

The invention relates to an insulin regulator construct comprising a glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene promoter and an insulin-sensitive element (ISE) of an insulin-like growth

CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 100.0%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 2.5e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
DB 1 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
QY 61 TTAAACCCGCTGCGGAGCGACCCCTTCAATAGGCCCTGGATAGGCCAGCATGTGT 120
DB 61 TTAAACCCGCTGCGGAGCGACCCCTTCAATAGGCCCTGGATAGGCCAGCATGTGT 120
QY 121 CCACCTGCCCGCGAGACAAACCCAGCAGCATTTGAACATGCAACAGGCCATCTGCC 180
DB 121 CCACCTGCCCGCGAGACAAACCCAGCAGCATTTGAACATGCAACAGGCCATCTGCC 180
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219

RESULT 2

ABX15373
ID ABX15373 standard; DNA; 270 BP.

AC ABX15373;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #1.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISF; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
OS Synthetic.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THUL P M.

XX Thule PM;

XX

DR WPI: 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

XX Claim 9; Page 13-14; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISF) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention

SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 100.0%; Score 219; DB 6; Length 270;
Best Local Similarity 100.0%; Pred. No. 2.7e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 60
DB 52 TCACAGCAAAACAACTTATTGTAACACGGGGATCTAGACGCTGCCCTGACATCA 111
QY 61 TTAAACCCGCTGCGGAGCGACCCCTTCAATAGGCCCTGGATAGGCCAGCATGTGT 120
DB 112 TTAAACCCGCTGCGGAGCGACCCCTTCAATAGGCCCTGGATAGGCCAGCATGTGT 171
QY 121 CCACCTGCCCGCGAGACAAACCCAGCAGCATTTGAACATGCAACAGGCCATCTGCC 180
DB 172 CCACCTGCCCGCGAGACAAACCCAGCAGCATTTGAACATGCAACAGGCCATCTGCC 231
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 219
DB 232 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCCGC 270

RESULT 3

ABX15374
ID ABX15374 standard; DNA; 321 BP.

AC ABX15374;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #2.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISF; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
OS Synthetic.

PN US2002107198-A1.

XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 219; DB 6; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.9e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTTAGACGCTGCTGACAAATCA 60
DB 103 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTTAGACGCTGCTGACAAATCA 162
QY 61 TTAAACCGTGTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 120
DB 163 TTAAACCGTGTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 222
QY 121 CCACTGCCCGCGAGACACAAACCCAGCGAGTTGAACATGTCACAGGCCATCTGGCCC 180
DB 223 CCACTGCCCGCGAGACACAAACCCAGCGAGTTGAACATGTCACAGGCCATCTGGCCC 282
QY 181 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 219
DB 283 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 321
RESULT 4
ABX15375
ID ABX15375 standard; DNA; 372 BP.
XX
XX ABX15375;
XX
XX 17-APR-2003 (first entry)
XX
XX Rat insulin regulator construct DNA #3.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;

KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS
OS Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 219; DB 6; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTTAGACGCTGCTGACAAATCA 60
DB 154 TCACAAGCAAAACAACTTATTGTAACAGGGGATCTTAGACGCTGCTGACAAATCA 213
QY 61 TTAAACCGTGTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 120
DB 214 TTAAACCGTGTGCGGAGCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGCATGGT 273
QY 121 CCACTGCCCGCGAGACACAAACCCAGCGAGTTGAACATGTCACAGGCCATCTGGCCC 180
DB 274 CCACTGCCCGCGAGACACAAACCCAGCGAGTTGAACATGTCACAGGCCATCTGGCCC 333
QY 181 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 219
DB 334 AGAGAGCTGTGACCACTTCGCTACTAGTAGCCGC 372

```

RESULT 5
ABK15376
ID ABK15376 standard; DNA; 423 BP.
XX
AC ABK15376;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #4.
XX
KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THULE/) THULE P M.
XX
PI Thule PW;
XX
DR MPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
Query Match 100.0%; Score 219; DB 6; Length 423;
Beet Local Similarity 100.0%; Pred. No. 3.1e-60;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGCAAAACAATTATTTGAACGCGGATCCAGACCGCTGCGCATCA 60
DB 205 TCACAGCAAAACAATTATTTGAACGCGGATCCAGACCGCTGCGCATCA 264
61 TTAACCGGTGCTGCGGACGACCGCTTCATAAGCCCTGGGTATGCGCAGCCAGCATGAT 120

```

```

DB 265 TTAACCGGTGCTGCGGACGACCGCTTCATAAGCCCTGGGTATGCGCAGCCAGCATGAT 324
QY 121 CCACGTCCCGCCGAGACCAAAACCAAGAGCATTTGAACACTGCACAGGCATGTGCC 180
DB 325 CCACGTCCCGCCGAGACCAAAACCAAGAGCATTTGAACACTGCACAGGCATGTGCC 384
QY 181 AGAGAGCTGTGACACACCACTTCGCTACTAGTACGCC 219
DB 385 AGAGAGCTGTGACACCACTTCGCTACTAGTACGCC 423
RESULT 6
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
XX
AC ABK63701;
XX
DT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX
KM Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KM differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
PN WO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US023872.
XX
PR 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244888P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
DR MPI; 2002-241625/29.
XX
CC Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
PS Claim 1; SEQ ID NO 1608; 239pp; English.
XX
CC The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
XX method can also be used to identify an agent which modulates the toxic
XX response and predict cellular pathways that a compound modulates in a
XX cell. The method utilises a set of at least two probes (on a solid
XX support in kit form), where each of the probes comprises a sequence that
XX specifically hybridises to a gene listed in the specification, a computer
XX system comprising a database containing information identifying the
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX set of genes comprising at least two genes listed in the specification,

```

CC and a user interface to view the information used to present information
CC identifying the expression level in a tissue or cell of at least one gene
CC listed in the specification. The method is useful for elucidating global
CC changes in gene expression and for identifying toxicity markers in
CC tissues or cell exposed to a known toxin. The genes may be used as
CC toxicity markers in drug screening and toxicity assays. The genes and
CC gene expression information may be used as diagnostic markers for the
CC prediction or identification of the physiological state of tissue or cell
CC sample that has been exposed to a compound or agent. Hepatotoxicity is
CC characterized by centrilobular necrosis and steatosis. The present
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC which is differentially expressed in response to a hepatotoxic agent
XX

SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCGAGGACATTGAACACTGCACACGGCCATCTGCCCGAGAGCTGTACAC 196
DB 8 CACAAACCCGAGGACATTGAACACTGCACACGGCCATCTGCCCGAGAGCTGTACAC 67

QY 197 CACTTCGGCTACTAGCTA 214
DB 68 CACTTCGGCTACTAGCTA 85

RESULT 7
ADBS2710
ID ADBS2710 standard; DNA; 1500 BP.
XX

AC ADBS2710;
XX

DT 04-DEC-2003 (first entry)
XX

DE Toxicity-related gene, SEQ ID 3227.
XX

KM Toxic; toxin; gene expression profile; hepatotoxicity; liver;
XX drug screening; toxicity assay; ds.
OS Unidentified.
XX

PN WO2003064624-A2.
XX

PD 07-AUG-2003.
XX

PF 31-JAN-2003; 2003WO-US003194.
XX

PR 31-JAN-2002; 2002US-00060087.
XX 15-MAR-2002; 2002US-0364045P.
XX 15-MAR-2002; 2002US-0364055P.
XX 30-DEC-2002; 2002US-0436643P.
XX

PA (GENE-) GENE LOGIC INC.
XX

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX

DR WPI: 2003-689530/65.
XX

PT Predicting a toxic effect of a compound, useful in identifying toxicity
XX markers in liver tissues or cells for drug screening and toxicity assays,
XX comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.
XX

PS Claim 1, SEQ ID NO 3227; 1156bp; English.
XX

CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a

CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 34.9%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCGAGGACATTGAACACTGCACACGGCCATCTGCCCGAGAGCTGTACAC 196
DB 8 CACAAACCCGAGGACATTGAACACTGCACACGGCCATCTGCCCGAGAGCTGTACAC 67

QY 197 CACTTCGGCTACTAGCTA 214
DB 68 CACTTCGGCTACTAGCTA 85

RESULT 8
ADBS2710
ID ADBS2710 standard; DNA; 1500 BP.
XX

AC ADBS2710;
XX

DT 04-DEC-2003 (first entry)
XX

DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX

KM toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
XX toxicity marker; toxicity progression; drug screening;
XX primary rat hepatocyte toxicity modelling; gene; ds.
OS Ratcus norvegicus.
XX

PN WO2003065993-A2.
XX

PD 14-AUG-2003.
XX

PF 04-FEB-2003; 2003WO-US003482.
XX

PR 04-FEB-2002; 2002US-0353171P.
XX 13-MAR-2002; 2002US-0363534P.
XX 08-APR-2002; 2002US-0370248P.
XX 10-APR-2002; 2002US-0371134P.
XX 10-APR-2002; 2002US-0371135P.
XX 10-APR-2002; 2002US-0371150P.
XX 11-APR-2002; 2002US-0371413P.
XX 19-APR-2002; 2002US-0373601P.
XX 19-APR-2002; 2002US-0373602P.
XX 22-APR-2002; 2002US-0374139P.
XX 08-MAY-2002; 2002US-0378370P.
XX 09-MAY-2002; 2002US-0378652P.
XX 09-MAY-2002; 2002US-0378653P.
XX 09-MAY-2002; 2002US-0378665P.
XX 09-JUL-2002; 2002US-0394230P.
XX 09-JUL-2002; 2002US-0394253P.
XX 04-SEP-2002; 2002US-0407688P.
XX 28-JAN-2003; 2003US-0442900P.
XX

PA (GENE-) GENE LOGIC INC.
XX

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Elashoff M;
XX

DR WPI; 2003-731472/69.

XX Determining if a compound induces a toxic effect on a tissue or cell, for

XX identifying hepatotoxic compounds, comprises comparing a gene expression

PT profile of a tissue or cell sample to a database of tox mean and non-tox

PT mean values.

XX Claim 44; SEQ ID NO 3252; 874bp; English.

XX The present invention describes a method for determining whether a

CC compound induces a toxic effect on a tissue or cell. The method comprises

CC preparing a gene expression profile of a tissue or cell sample exposed to

CC the compound, and comparing the gene expression profile to a database

CC comprising data or information on the Tox mean and non-Tox mean value.

CC The method is useful for predicting or identifying at least one toxic

CC effect, particularly hepatotoxicity, of a test or unknown compound. The

CC genes listed in the specification are useful as diagnostic or toxicity

CC markers for the prediction or identification of the physiological state

CC of tissue or cell sample that has been exposed to a compound, or to

CC identify or predict the toxic effects of a compound or an agent. These

CC may also be used as markers for monitoring toxicity progression or for

CC drug screening. The present sequence represents a primary rat hepatocyte

CC toxicity modelling related gene sequence from the present invention.

SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

XX

Query Match 34.9%; Score 76.4; DB 10; Length 1500;

Best Local Similarity 98.7%; Pred. No. 3.6e-14;

Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCGAGGAGATTGAACTGACACGCGCATCTGCCCAGAGAGCTGTGACAC 196

DB 8 CACAAACCCGAGGAGATTGAACTGACACGCGCATCTGCCCAGAGAGCTGTGACAC 67

QY 197 CACTTCGGCTACTAGCTA 214

DB 68 CACTTCGGCTACTAGCTA 85

RESULT 9

ID ABT41911

XX ABT41911 standard; DNA; 1500 BP.

AC ABT41911;

XX

DT 26-JUN-2003 (first entry)

XX

DE Toxicity modelling related rat gene SEQ ID No 1613.

XX

KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;

KW database; drug screening; toxicity assay; rat; ds.

OS

XX Rattus norvegicus.

XX

PN WO200295000-A2.

XX

PD 28-NOV-2002.

XX

PF 22-MAY-2002; 2002WO-US016173.

XX

XX 22-MAY-2001; 2001US-0292335P.

XX 13-JUN-2001; 2001US-0297523P.

PR 19-JUN-2001; 2001US-0298925P.

PR 10-JUL-2001; 2001US-0303807P.

PR 10-JUL-2001; 2001US-0303808P.

PR 10-JUL-2001; 2001US-0303810P.

PR 28-AUG-2001; 2001US-0315047P.

PR 27-SEP-2001; 2001US-0324928P.

PR 22-OCT-2001; 2001US-0330462P.

PR 01-NOV-2001; 2001US-0330867P.

PR 21-NOV-2001; 2001US-0331805P.

PR 06-DEC-2001; 2001US-0336144P.

PR 19-DEC-2001; 2001US-0340873P.

XX

PR 21-FEB-2002; 2002US-0357842P.

PR 21-FEB-2002; 2002US-0357843P.

PR 21-FEB-2002; 2002US-0357844P.

PR 15-MAR-2002; 2002US-0364134P.

PR 08-APR-2002; 2002US-0370144P.

PR 08-APR-2002; 2002US-0370206P.

PR 08-APR-2002; 2002US-0370247P.

PR 17-APR-2002; 2002US-0372794P.

PR 21-APR-2002; 2002US-0371679P.

XX

PA (GENE-) GENE LOGIC INC.

XX

PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;

XX

DR WPI; 2003-148464/14.

XX

PT Predicting at least one toxic effect of a compound, useful for toxicity

PT modeling, comprises preparing a gene expression profile of a tissue or

PT cell sample exposed to the compound, and comparing the gene expression

PT profile to a database.

XX

PS Example 4; Page; 446pp; English.

XX

CC The invention relates to a novel method of predicting at least one toxic

CC effect of a compound. The method comprises a gene expression profile of a

CC tissue or cell sample exposed to the compound, and comparing the gene

CC expression profile to a database comprising at least part of the data or

CC information given in the specification. The methods are useful for

CC predicting at least one toxic effect of a compound, predicting the renal

CC progression of a toxic effect of a compound, predicting the renal

CC toxicity of a compound, or identifying toxicity markers in tissues or

CC cells exposed to known renal toxin. The genes are useful as toxicity

CC markers in drug screening and toxicity assays, in monitoring disease or

CC physiological states, or disease progression. This polynucleotide

CC represents a rat DNA sequence relating to the toxic effect database

CC described in the specification. NOTE: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the World Intellectual Property

CC Organization

XX

SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

XX

Query Match 34.9%; Score 76.4; DB 10; Length 1500;

Best Local Similarity 98.7%; Pred. No. 3.6e-14;

Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCGAGGAGATTGAACTGACACGCGCATCTGCCCAGAGAGCTGTGACAC 196

DB 8 CACAAACCCGAGGAGATTGAACTGACACGCGCATCTGCCCAGAGAGCTGTGACAC 67

QY 197 CACTTCGGCTACTAGCTA 214

DB 68 CACTTCGGCTACTAGCTA 85

RESULT 10

ID ADP72689

XX ADP72689 standard; DNA; 1500 BP.

XX

AC ADP72689;

XX

DT 26-AUG-2004 (first entry)

XX

DE Renal toxin progression gene marker #1278.

XX

KW ds; toxic effect; gene expression profile; kidney tissue;

KW differential gene expression; toxicity progression; toxicity marker;

KW drug screening; toxicity assay; kidney pathology; nephritis;

KW kidney necrosis; glomerular injury; tubular injury;

KW focal segmental glomerulosclerosis.

OS

XX Rattus norvegicus.

XX

PN WO2004048598-A2.
XX
PD 10-JUN-2004.
XX
XX 24-NOV-2003; 2003WO-US037556.
XX PF
XX 22-NOV-2002; 2002US-00301856.
XX PR
XX (GENE-) GENE LOGIC INC.
XX PA
XX Mendick DL, Porter MM, Johnson KR, Castle A, Higge B;
XX PI Elashoff M;
XX PI
XX WPI; 2004-460771/43.
XX DR
XX
XX Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
XX Claim 11; SEQ ID NO 1278; 266pp; English.
XX
XX The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 34.9%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.6e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 137 CACAAACCCAGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 196
DB 8 CACAAACCCAGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 67
QY 197 CACTTCGGCTACTAGCTA 214
DB 68 CACTTCGGCTACTAGCTA 85
RESULT 11
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
XX AAH22429;
XX
XX 22-AUG-2001 (first entry)
XX DT
XX Rat insulin-like growth factor binding protein nucleotide sequence.
XX DE
XX
XX Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KW non-steroidal antiinflammatory drug; ds.

XX
OS Rattus norvegicus.
XX
XX WO200138579-A2.
XX PN
XX 31-MAY-2001.
XX PD
XX 21-NOV-2000; 2000WO-US032049.
XX PF
XX 22-NOV-1999; 99US-0166923P.
XX PR 18-FEB-2000; 2000US-0183531P.
XX PR 20-NOV-2000; 2000US-00717321.
XX
XX (CURA-) CURAGEN CORP.
XX PA
XX Gould-Rothberg BE, DiPippo VA, Ramseh TM, Gerwein RW;
XX PI
XX WPI; 2001-355948/37.
XX DR
XX
XX Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
XX Disclosure; Page 22-24; 76pp; English.
XX
XX The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in a reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;
Query Match 34.9%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 5.3e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 137 CACAAACCCAGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 196
DB 1 CACAAACCCAGAGCATTTGAACACTGCACAGCGGCATCTGCCAGAGAGCTGTGACCAC 60
QY 197 CACTTCGGCTACTAGCTA 214
DB 61 CACTTCGGCTACTAGCTA 78
RESULT 12
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
XX ABN95896;
XX
XX 13-AUG-2002 (first entry)
XX DT
XX Gene #2394 used to diagnose liver cancer.
XX DE
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytotoxic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
XX OS

```
PN WO200229103-A2.
XX
XX 11-APR-2002.
PD
XX 02-OCT-2001; 2001WO-US030589.
XX
XX 02-OCT-2000; 2000US-0237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG,
XX WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer, hepatocellular
XX carcinoma or metastatic liver tumor in a patient, involves detecting the
XX level of expression of two or more genes in a liver tissue sample.
XX
XX Claim 1; SEQ ID NO 2394; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
XX
XX Query Match 27.8%; Score 60.8; DB 6; Length 6128;
XX Best Local Similarity 66.5%; Pred. No. 5.9e-09;
XX Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
XX
XX 2 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCGCTGACATCAT 61
XX 473 CACTAGCAAAACAACTTATTTTGAACACTGCTCTAGCGTGCGGCGCTGCCATCAT 532
XX
XX 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 111
XX 533 TAACCTCTGCTGTGCAAGTGCGCGGCTGTGCTTTTAAAGTGCGCGCTGTGTCCAGC 592
XX
XX 112 CAGCATGTGTCAGTGGCGCGCGAGACACAAACCCGAGCATTTGAACACTGTC-ACACGG 170
XX 593 GAGCATTCGGCCACCGCCATCC-----CATTCAGCGAGCATCTGCGCGCGCGCGCG 644
XX
XX 171 CCATCTGCCAGAGAGCTGTGACCAACCACTTC 202
XX 645 CCACCTCCCGACAGAGCACTGGCCACCGCTCC 676
XX
XX
XX RESULT 13
XX ABV75371
XX ID ABV75371 standard; DNA; 6128 BP.
XX
XX AC ABV75371;
XX
XX 07-MAR-2003 (first entry)
XX
XX Human IGFBP-1 gene sequence.
XX
XX Insulin-like growth factor binding protein; IGFBP; cyrostatic; liver;
XX cancer; human; IGFBP-1; gene; ds.
XX
XX Homo sapiens.
XX
XX OS
```

```
XX Key Location/Qualifiers
XX FH 762..5200
XX CDS /tag=.5200
XX
XX /tag=a
XX FT /product= "IGFBP-1"
XX FT /note= "insulin-like growth factor binding protein;
XX FT contains introns"
XX FT 762..1110
XX FT /tag=b
XX FT /number=1
XX FT 1111..2656
XX FT /tag=c
XX FT /number=1
XX FT 2657..2826
XX FT /tag=d
XX FT /number=2
XX FT 2827..4040
XX FT /tag=e
XX FT /number=2
XX FT 4041..4169
XX FT /tag=f
XX FT /number=3
XX FT 4170..5068
XX FT /tag=g
XX FT /number=3
XX FT 5069..5197
XX FT /tag=h
XX FT /number=4
XX
XX WO200290580-A1.
XX
XX 14-NOV-2002.
XX
XX 03-MAY-2002; 2002WO-AU000558.
XX
XX 03-MAY-2001; 2001US-0288441P.
XX
XX (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
XX PA (SIGE-) SINGAPORE GEN HOSPITAL PTE LTD.
XX PA (ARGA/) ARGAE V P.
XX
XX Huynh TH, Chow PKH, Soo KC;
XX WPI; 2003-103522/09.
XX DR P-PSDB; ABB82757.
XX
XX Detecting the presence or diagnosing the risk of a liver cancer in a
XX patient comprises detecting aberrant expression of a gene encoding an
XX insulin-like growth factor binding protein.
XX
XX Example; Page 104-108; 142bp; English.
XX
XX The invention relates to detecting the presence or diagnosing the risk of
XX a liver cancer in a patient. The method involves detecting in a
XX biological sample obtained from the patient aberrant expression of a gene
XX encoding an insulin-like growth factor binding protein (IGFBP). The
XX method is useful for detecting the presence or diagnosing the risk of a
XX liver cancer or for screening agents in a patient. The agent is useful
XX for the manufacture of a medicament for treating and/or preventing liver
XX cancer. The present sequence represents a human IGFBP-1 polypeptide
XX encoding genomic DNA (GenBank Accession No. M74587)
XX
XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
XX
XX Query Match 27.8%; Score 60.8; DB 10; Length 6128;
XX Best Local Similarity 66.5%; Pred. No. 5.9e-09;
XX Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
XX
XX 2 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCGCTGACATCAT 61
XX 473 CACTAGCAAAACAACTTATTTTGAACACTGCTCTAGCGTGCGGCGCTGCCATCAT 532
XX
XX 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 111
```


Db	2185	CACTAACAAACAATACTTATTAACTCAACTCTCTACGACGACGCTACCAATCAT	2126
Qy	62	TAAAC	66
Db	2125	TAAAC	2121

Search completed: September 1, 2005, 00:34:41
Job time : 238.225 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 16.1069 seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51

Sequence: 1 catggcgagcgagcgagcgagcgc.....ggactctggcccccagtgta 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	194	US-08-945-140-1	Sequence 1, Appl
2	51	100.0	13011	US-08-791-849A-14	Sequence 14, Appl
3	36.6	71.8	15108	US-09-949-016-11786	Sequence 11786, A
4	36.6	71.8	15108	US-09-949-016-17205	Sequence 17205, A
5	24.4	47.8	3138	US-09-234-332-5	Sequence 5, Appl
6	24	47.1	24	US-08-945-140-2	Sequence 2, Appl
7	24	47.1	1498	US-08-775-428-1	Sequence 1, Appl
8	23.4	45.9	601	US-09-949-016-60944	Sequence 60944, A
9	23.4	45.9	659	US-09-016-434-1017	Sequence 1017, Ap
10	23.4	45.9	2777	US-09-282-147-39	Sequence 39, Appl
11	23.4	45.9	3078	US-09-949-016-11956	Sequence 11956, Ap
12	23.4	45.9	3088	US-09-023-655-1294	Sequence 1294, Ap
13	23.4	45.9	46244	US-09-949-016-13508	Sequence 13508, A
14	23.2	45.5	601	US-09-949-016-49730	Sequence 49730, A
15	23.2	45.5	601	US-09-949-016-49731	Sequence 49731, A
16	23.2	45.5	601	US-09-949-016-49732	Sequence 49732, A
17	23.2	45.5	126200	US-09-949-016-11824	Sequence 11824, A
18	23.2	45.5	126200	US-09-949-016-13193	Sequence 13193, A
19	23	45.1	455	US-09-949-016-60945	Sequence 60945, A
20	23	45.1	455	US-09-949-016-17115	Sequence 17115, A
21	22.8	44.7	118067	US-09-949-016-12208	Sequence 12208, A
22	22.6	44.7	118067	US-09-949-016-12208	Sequence 12208, A
23	22.4	43.9	30244	US-09-949-016-13550	Sequence 13550, A
24	22.4	43.9	30244	US-09-949-016-12742	Sequence 12742, A
25	22.4	43.9	97424	US-09-949-016-15576	Sequence 15576, A
26	22.4	43.9	97424	US-09-949-016-15576	Sequence 15576, A
27	22.2	43.5	1001	US-09-023-655-805	Sequence 805, Appl

28	22	43.1	1257	4	US-09-826-509-452	Sequence 452, Appl
29	22	43.1	1472	4	US-09-016-434-1388	Sequence 1388, Ap
30	22	43.1	41522	4	US-09-949-016-11932	Sequence 11932, A
31	22	43.1	41523	4	US-09-949-016-15764	Sequence 15764, A
32	22	43.1	26465	4	US-09-949-016-13747	Sequence 13747, A
33	21.8	42.7	601	4	US-09-949-016-116783	Sequence 116783, A
34	21.8	42.7	798	4	US-09-252-991A-11759	Sequence 11759, A
35	21.8	42.7	1314	4	US-09-252-991A-11608	Sequence 11608, A
36	21.8	42.7	1716	4	US-09-902-540-2799	Sequence 2799, Ap
37	21.8	42.7	4236	4	US-09-252-991A-11837	Sequence 11837, A
38	21.8	42.7	14570	4	US-09-902-540-1012	Sequence 1012, Ap
39	21.8	42.7	325034	4	US-09-949-016-14957	Sequence 14957, A
40	21.8	42.7	389504	4	US-09-949-016-11774	Sequence 11774, A
41	21.6	42.4	601	4	US-09-949-016-148899	Sequence 148899, A
42	21.6	42.4	601	4	US-09-949-016-148900	Sequence 148900, A
43	21.6	42.4	993	4	US-09-949-016-3977	Sequence 3977, Appl
44	21.6	42.4	1980	4	US-09-697-022-1	Sequence 1, Appl
45	21.6	42.4	8392	1	US-08-080-255-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-945-140-1
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky & Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
SYNDROMES: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-945-140-1
Query Match 100.0%; Score 51; DB 3; Length 194;

Best Local Similarity 100.0%; Pred. No. 6.3e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 11 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 61

RESULT 2

US-08-791-849A-14
; Sequence 14, Application US/08791849A
; Patent No. 5914449
; GENERAL INFORMATION:
; APPLICANT: Makoto MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,849A
; FILING DATE: January 30, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: rat (Rattus norvegicus)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(3212..3218, 3766..3948, 5917..6008,
; 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; LOCATION: 9298..9479, 10163..10269)
; FEATURE:
; NAME/KEY: exon
; LOCATION: join(3194..3218, 3766..3948, 5917..6008,
; 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; FEATURE:
; NAME/KEY: intron
; LOCATION: join(3219..3765, 3949..5916, 6009..6151,
; 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
; LOCATION: 9480..10162)
US-08-791-849A-14

Query Match 100.0%; Score 51; DB 2; Length 13011;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 3021 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 3071

RESULT 3

US-09-949-016-11786
; Sequence 11786, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11786
; LENGTH: 15108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11786

Query Match 71.8%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.0012;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 2357 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 2399

RESULT 4

US-09-949-016-17205
; Sequence 17205, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 17205
; LENGTH: 15108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17205

Query Match 71.8%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.0012;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 51
|||||
DB 2357 CACGGGCGACTCCCGTGTCTCTGACTCTGGCCCCCAGTGA 2399

RESULT 5
US-09-234-332-5
Sequence 5, Application US/09234332A
Patent No. 6087168
GENERAL INFORMATION:
APPLICANT: Cedars-Sinai Medical Center
APPLICANT: Michel F. Levesque, M.D.
APPLICANT: Thomas Neuman, Ph.D.
TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
FILE REFERENCE: P07 41494
CURRENT APPLICATION NUMBER: US/09/234,332A
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 3138
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: gene
LOCATION: (0)...(0)
OTHER INFORMATION: z1c 1 Protein gene; Genbank Accession D76435
US-09-234-332-5

Query Match 47.8%; Score 24.4; DB 3; Length 3138;
Best Local Similarity 73.8%; Pred. No. 20;
Matches 31; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 7 GCGACGGGACATCCCGTGGTCTCTGGACTCTGGCCCCCAGT 48
DB 764 GCGCCGACGACGACGATGCTCTGGACGCGGCCCCCAGT 805

RESULT 6
US-08-945-140-2
Sequence 2, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: 500 Arcola Road, Mallescop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
US-08-945-140-2

Query Match 47.1%; Score 24; DB 3; Length 24;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GCGACGGGACATCCCGTGGTTC 29
DB 1 GCGACGGGACATCCCGTGGTTC 24

RESULT 7
US-08-775-428-1/c
Sequence 1, Application US/08775428
Patent No. 5976834
GENERAL INFORMATION:
APPLICANT: Sathe, Ganesh
APPLICANT: Fueterer, Wendy
APPLICANT: Bergema, Derek
APPLICANT: Ellis, Catherine
TITLE OF INVENTION: CDNA CLONE HNF1D15 THAT ENCODES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,428
FILING DATE: 09-JAN-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50042
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4060
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1498 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-775-428-1

Query Match 47.1%; Score 24; DB 2; Length 1498;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 8 GCACGGGCGACTCCGTGTTCTTGACTCTGGCCCCCAG 47
Db 1135 GCATGTGACACACCCCGGAGCTGTGACTCTGGCATTCCAG 1096

RESULT 8
US-09-949-016-60944
Sequence 60944, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 60944
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-60944

Query Match
Best Local Similarity 45.9%; Score 23.4; DB 4; Length 601;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 CATGGCGCAGCGGGGACTCCGTGTTCTTGACTCTGGCCCCCAGTG 49
Db 61 CCTGGGCTCAGGGGCGACCTCTCTGCTGCTCTGCTGGGCCGCTGTG 109

RESULT 9
US-09-016-434-1017/c
Sequence 1017, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSES: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1017:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CERVN0701
CLONE: 930839
US-09-016-434-1017

Query Match
Best Local Similarity 45.9%; Score 23.4; DB 4; Length 659;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 11 CGGGGCACTCCCGTGTCTTGACTCTGGCCCCCAGTGTA 51
Db 388 CGAGCACTGACGAGAGCTGTGACTCTGTGCCCATGTCA 348

RESULT 10
US-09-282-147-39/c
Sequence 39, Application US/09282147
Patent No. 6274147
GENERAL INFORMATION:
APPLICANT: YAO, Kun
TITLE OF INVENTION: METHOD FOR GENERATING NONPATHOGENIC, INFECTIOUS
TITLE OF INVENTION: PANCREATIC NECROSIS VIRUS (IPNV) FROM SYNTHETIC RNA
FILE REFERENCE: 8288-9023
CURRENT APPLICATION NUMBER: US/09/282,147
CURRENT FILING DATE: 1999-03-31
EARLIER APPLICATION NUMBER: US/60/080,278
EARLIER FILING DATE: 1998-03-31
EARLIER APPLICATION NUMBER: PCT/US97/12955
EARLIER FILING DATE: 1998-03-31
NUMBER OF SEQ ID NOS: 51
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 39
LENGTH: 2777
TYPE: DNA
ORGANISM: Infectious pancreatic necrosis virus
US-09-282-147-39

Query Match
Best Local Similarity 45.9%; Score 23.4; DB 3; Length 2777;
Matches 30; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 10 ACGGGGCACTCCCGTGTCTTGACTCTGGCCCCCAGTGT 50
Db 1077 AGGGGGCACTCCAGATGTCTCTGCTTCAAGTCCCACTCT 1037

RESULT 11
US-09-949-016-1956/c
Sequence 1956, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498


```
; Sequence 49731, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49731
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-49731

Query Match      45.5%; Score 23.2; DB 4; Length 601;
Best Local Similarity 70.5%; Pred. No. 44;
Matches 31; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      4 GGGCGACGGGGGACTCCCGTGTCTCTGAGACTCTGGCCCCCAG 47
         |||||
Db      452 GAGCCAACCGCGCCCTCCCTGTGATCTTGAGCTCCAGCCTCCAG 409
```

Search completed: September 1, 2005, 07:07:03
Job time : 22.2736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 31, 2005, 12:52:03 ; Search time 286.304 Seconds
(without alignments)
5582.628 Million cell updates/sec

Title: US-09-972-916B-3

Perfect score: 270

Sequence: 1 catggcgcgccggcgccaccc.....ttccgctactagctagcgcg 270

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980a:*
2: geneseqn1980b:*
3: geneseqn2000a:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	100.0	270	6	ABX15373 Rat insul
2	270	100.0	423	6	ABX15376 Rat insul
3	219.4	81.3	321	6	ABX15374 Rat insul
4	219.4	81.3	372	6	ABX15375 Rat insul
5	219	81.1	219	6	ABX15372 Rat insul
6	76.4	28.3	1500	6	ABK63701 Rat segue
7	76.4	28.3	1500	10	ADBS8201 ADBS8201
8	76.4	28.3	1500	10	ADBS2710 ADBS2710
9	76.4	28.3	1500	10	ABR41911 ABR41911
10	76.4	28.3	1500	12	ADP72689 ADP72689
11	76.4	28.3	5001	4	AAH22429 AAH22429
12	60.8	22.5	6128	6	ABN95896 ABN95896
13	60.8	22.5	6128	10	ABV75371 ABV75371
14	60.8	22.5	6134	4	AAH57489 AAH57489
15	51	18.9	51	6	ABX15371 Rat glucoc
16	51	18.9	194	2	AAH73002 AAH73002
17	51	18.9	321	6	ABX15374 Rat insul
18	51	18.9	372	6	ABX15375 Rat insul
19	51	18.9	13011	2	AAH96631 AAH96631
20	51	18.9	13011	10	ABH42448 ABH42448

21	51	18.9	13011	12	ADP72914 ADP72914
22	48	17.8	48	6	ABX15379 Rat liver
23	48	17.8	51	6	ABX15380 Rat liver
24	37.8	14.0	7061	6	ABX152997 Human imm
25	37.8	14.0	7061	6	ABX170248 Chemical1
26	37.8	14.0	7061	6	AAH61189 Human gen
27	35	13.0	3895	6	ABN87725 Human pro
28	33.4	12.4	1806	6	ABO75795 Human MDD
29	33.4	12.4	2667	11	ADM03554 ADM03554
30	33.2	12.3	650	8	ABH52705 ABH52705
31	32.4	12.0	1048	5	AAH69850 AAH69850
32	32.2	11.9	4447	2	AAH14078 AAH14078
33	32	11.9	76180	13	ABD33385 ABD33385
34	31.6	11.7	548	13	ADQ51481 ADQ51481
35	31.6	11.7	927	6	ABH57730 ABH57730
36	31.6	11.7	125	6	AAH94440 AAH94440
37	31.6	11.7	1129	6	ADH33646 ADH33646
38	31.6	11.7	1288	10	ADC87210 ADC87210
39	31.6	11.7	1344	6	ABH57731 ABH57731
40	31.6	11.7	1348	8	ABX71176 ABX71176
41	31.6	11.7	1513	6	ADH16493 ADH16493
42	31.6	11.7	1550	4	AAH94450 AAH94450
43	31.6	11.7	1555	4	AAH90027 AAH90027
44	31.6	11.7	1555	5	AAH39684 AAH39684
45	31.6	11.7	1555	9	ADB32644 ADB32644

ALIGNMENTS

RESULT 1	ABX15373
ID	ABX15373 standard; DNA; 270 BP.
AC	ABX15373;
XX	
DT	17-APR-2003 (first entry)
XX	
DE	Rat insulin regulator construct DNA #1.
KW	Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGF1R-1;
KW	insulin-sensitive element; ISE; basal promoter; hyperglycemia; insulin;
KW	insulin-like growth factor binding protein-1; hypoglycemia; glucose;
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW	glucagon; euglycemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW	intracellular hormone receptor; insulin regulator construct; anabolic.
XX	
OS	Rattus norvegicus.
OS	Synthetic.
PN	US2002107198-A1.
XX	
PD	08-AUG-2002.
XX	
PF	10-OCT-2001; 2001US-00972916.
XX	
PR	11-OCT-2000; 2000US-0239113P.
XX	
PA	(THUL/) THUL P M.
XX	
FI	Thule PM;
XX	
DR	WPI; 2002-674190/72.
XX	
PT	New insulin regulator cassette, useful e.g. for treating diabetes,
XX	provides specific, glucose-inducible transgenic expression of insulin in
PT	liver cells.
XX	
PS	Claim 9; Page 13-14; 37pp; English.
XX	
CC	The invention relates to an insulin regulator construct comprising a
XX	glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 100.0%; Score 270; DB 6; Length 270;
 Best Local Similarity 100.0%; Pred. No. 3.2e-77;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 DB 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 QY 61 AAACAACCTTATTTTGAACACGGGGATCTTACGACCGCTGCCATCATTAACCGGT 120
 DB 61 AAACAACCTTATTTTGAACACGGGGATCTTACGACCGCTGCCATCATTAACCGGT 120
 QY 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGCTCCAGTGGCC 180
 DB 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGCTCCAGTGGCC 180
 QY 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACGCGCATCTGCCAGAGAGCTG 240
 DB 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACGCGCATCTGCCAGAGAGCTG 240
 QY 241 TGACCAACCACTTCGGCTACTAGTACCGCG 270
 DB 241 TGACCAACCACTTCGGCTACTAGTACCGCG 270

RESULT 2

ABX15376
 ID ABX15376 standard; DNA; 423 BP.

AC ABX15376;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #4.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 XX intracellular hormone receptor; insulin regulator construct; anabolic.

OS Ratus norvegicus.
 OS Synthetic.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THULE P M.
 PA
 XX Thule PM;
 PT WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.

XX Claim 9; Page 14; 37pp; English.

XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match 100.0%; Score 270; DB 6; Length 423;
 Best Local Similarity 100.0%; Pred. No. 3.8e-77;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 60
 DB 154 CATGGCGCAGCGGGGCACTCCCGTGGTCTGTGACTCTGAGCCGCCAGTGTATCACAAGCA 213
 QY 61 AAACAACCTTATTTTGAACACGGGGATCTTACGACCGCTGCCATCATTAACCGGT 120
 DB 214 AAACAACCTTATTTTGAACACGGGGATCTTACGACCGCTGCCATCATTAACCGGT 273
 QY 121 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGCTCCAGTGGCC 180
 DB 274 GCTGCCGAGCCAGCGCCCTTATTAAGGCGCTGGGTATGGCCAGCAGATGCTCCAGTGGCC 333
 QY 181 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACGCGCATCTGCCAGAGAGCTG 240
 DB 334 GCCGAGACACAAACCCAGCAGCATTTGAACACTGTACACGCGCATCTGCCAGAGAGCTG 393
 QY 241 TGACCAACCACTTCGGCTACTAGTACCGCG 270
 DB 394 TGACCAACCACTTCGGCTACTAGTACCGCG 423

RESULT 3

ABX15374
 ID ABX15374 standard; DNA; 321 BP.

AC ABX15374;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #2.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;

KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
OS
XX Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
SQ
Query Match 81.3%; Score 219.4; DB 6; Length 321;
Best Local Similarity 99.5%; Pred. No. 8.7e-61;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 50 TATCAAGCAAAACAACCTATTGAAACGGGGATCCTTAGACAGCTGCCCTGACAT 109
DB 101 TGTCAAGCAAAACAACCTATTGAAACGGGGATCCTTAGACAGCTGCCCTGACAT 160
QY 110 CATTAAACCGGTGCTGCCAGACGAGCCCTTCATTAAGGCGCTGGGTATGGCCAGCCAGCATG 169
DB 161 CATTAAACCGGTGCTGCCAGACGAGCCCTTCATTAAGGCGCTGGGTATGGCCAGCCAGCATG 220
QY 170 GTTCACTGCCCGCCGAGACCAAAACCCAGAGCATTTGAACACTGCACAGGCCCATCTGC 229
DB 221 GTTCACTGCCCGCCGAGACCAAAACCCAGAGCATTTGAACACTGCACAGGCCCATCTGC 280
QY 230 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTGACGCGC 270
DB 281 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTGACGCGC 321
RESULT 4
ABX15375

ID ABX15375 standard; DNA; 372 BP.
XX
XX AC ABX15375;
XX
XX 17-APR-2003 (first entry)
XX
XX DE Rat insulin regulator construct DNA #3.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS
OS Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
SQ
Query Match 81.3%; Score 219.4; DB 6; Length 372;
Best Local Similarity 99.5%; Pred. No. 9.2e-61;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 50 TATCAAGCAAAACAACCTATTGAAACGGGGATCCTTAGACAGCTGCCCTGACAT 109
DB 152 TGTCAAGCAAAACAACCTATTGAAACGGGGATCCTTAGACAGCTGCCCTGACAT 211
QY 110 CATTAAACCGGTGCTGCCAGACGAGCCCTTCATTAAGGCGCTGGGTATGGCCAGCCAGCATG 169
DB 212 CATTAAACCGGTGCTGCCAGACGAGCCCTTCATTAAGGCGCTGGGTATGGCCAGCCAGCATG 271

```
QY 170 GTCCATGCCCCGCGAGACAAACCAGAGCATTTGAACACTGACACGGCCATCTGC 229
DB 272 GTCCATGCCCCGCGAGACAAACCAGAGCATTTGAACACTGACACGGCCATCTGC 331
QY 230 CCAGAGAGCTGTGACCAACCACTTCCGCTACTAGTAGCCGC 270
DB 332 CCAGAGAGCTGTGACCAACCACTTCCGCTACTAGTAGCCGC 372

RESULT 5
ABX15372
ID ABX15372 standard; DNA; 219 BP.
AC ABX15372;
XX
XX 17-APR-2003 (first entry)
XX
XX Rat insulin-sensitive element (ISE) DNA.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
XX
XX OS
XX US2002107198-A1.
XX
XX PD 08-AUG-2002.
XX
XX PF 10-OCT-2001; 2001US-00972916.
XX
XX PR 11-OCT-2000; 2000US-0239113P.
XX
XX PA (THUL/) THULE P M.
XX
XX PI Thule PM;
XX
XX DR WPI; 2002-674190/72.
XX
XX PT New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX PS Claim 8; Page 13; 37pp; English.
XX
XX CC The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin-sensitive element of the invention
XX
XX SO Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 81.1%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 1e-60;
```

```
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 52 TCACAGAGAAAACAACTTATTTTGAACACGGGATCTTACACGCTCCCTGACAAATCA 111
DB 1 TCACAGAGAAAACAACTTATTTTGAACACGGGATCTTACACGCTCCCTGACAAATCA 60
QY 112 TTAACCCGTGCTGCGAGACCAAGCCCTTCATTAAGCCCTGGGTATAGCCAGCAGATGAT 171
DB 61 TTAACCCGTGCTGCGAGACCAAGCCCTTCATTAAGCCCTGGGTATAGCCAGCAGATGAT 120
QY 172 CCACGTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGACACGGCCATCTGCCC 231
DB 121 CCACGTGCCCCGCGAGACAAACCAGAGCATTTGAACACTGACACGGCCATCTGCCC 180
QY 232 AGAGAGCTGTGACCAACCACTTCCGCTACTAGTAGCCGC 270
DB 181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGTAGCCGC 219

RESULT 6
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
XX
XX ABK63701;
XX
XX AC
XX 18-JUN-2002 (first entry)
XX
XX DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX
XX KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
XX OS Rattus norvegicus.
XX
XX PN WO200210453-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 30-JUL-2001; 2001WO-US023872.
XX
XX PR 31-JUL-2000; 2000US-0222040P.
XX PR 02-NOV-2000; 2000US-0244880P.
XX PR 11-MAY-2001; 2001US-0290029P.
XX PR 15-MAY-2001; 2001US-0290645P.
XX PR 22-MAY-2001; 2001US-0292336P.
XX PR 06-JUN-2001; 2001US-0295798P.
XX PR 13-JUN-2001; 2001US-0297457P.
XX PR 19-JUN-2001; 2001US-0298884P.
XX PR 09-JUL-2001; 2001US-0303459P.
XX
XX PA (GENE-) GENE LOGIC INC.
XX
XX PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
XX DR WPI; 2002-241625/29.
XX
XX PT Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
XX PS Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX CC The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
```

CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilize a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridizes to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity is
 CC characterized by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent
 CC
 XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 6; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 247
 DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 68 CACTTCGGCTACTAGCTA 85

RESULT 7
 ADB58201
 ID ADB58201 standard; DNA; 1500 BP.

AC ADB58201;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3227.
 XX
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 KW drug screening; toxicity assay; da.
 XX
 OS Unidentified.
 XX
 PN WO2003064624-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 31-JAN-2003; 2003WO-US003194.
 XX
 PR 31-JAN-2002; 2002US-00060087.
 PR 15-MAR-2002; 2002US-0364045P.
 PR 15-MAR-2002; 2002US-0364055P.
 PR 30-DEC-2002; 2002US-0436643P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 XX
 DR WPI; 2003-689530/65.
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 PT markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 PT the compound.
 XX

PS Claim 1; SEQ ID NO 3227; 1156bp; English.
 XX
 CC The present invention relates to a method for predicting a toxic effect
 CC of a compound. The method comprises preparing a gene expression profile
 CC of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 CC of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

Seq Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 28.3%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 247
 DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 68 CACTTCGGCTACTAGCTA 85

RESULT 8
 ADB52710
 ID ADB52710 standard; DNA; 1500 BP.

AC ADB52710;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3852.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; da.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003065993-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003WO-US003482.
 XX
 PR 04-FEB-2002; 2002US-0353171P.
 PR 13-MAR-2002; 2002US-0365344P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.
 PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373602P.
 PR 19-APR-2002; 2002US-0374139P.
 PR 22-APR-2002; 2002US-0378370P.
 PR 08-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-JUL-2002; 2002US-0394230P.
 PR 09-JUL-2002; 2002US-0394253P.

```
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
PA
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX
PS Claim 44; SEQ ID NO 3252; 874pp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.4e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCGAGGAGCATTTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACCC 247
DB 8 CACAAACCGAGGAGCATTTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACCC 67

QY 248 CACTTCGCTACTAGCTA 265
DB 68 CACTTCGCTACTACTA 85

RESULT 9
ABP71911
ID ABP71911 standard; DNA; 1500 BP.
XX
XX ABP71911;
XX
XX 26-JUN-2003 (first entry)
XX
XX Toxicity modelling related rat gene SEQ ID No 1613.
XX
XX Toxic effect; gene expression profile; renal toxicity; toxicity marker;
XX database; drug screening; toxicity assay; rat; ds.
XX
XX Ratcus norvegicus.
XX
XX
XX WO200295000-A2.
XX
XX
XX 28-NOV-2002.
XX
XX
XX 22-MAY-2002; 2002WO-US016173.
XX
XX
XX 22-MAY-2001; 2001US-0292335P.
XX
XX 13-JUN-2001; 2001US-0297523P.
XX
XX 19-JUN-2001; 2001US-0298925P.
XX
XX 10-JUL-2001; 2001US-0303807P.
XX
XX 10-JUL-2001; 2001US-0303808P.
XX
```

```
PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-148464/14.
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
XX modeling, comprises preparing a gene expression profile of a tissue or
XX cell sample exposed to the compound, and comparing the gene expression
XX profile to a database.
XX
XX Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
XX effect of a compound. The method comprises a gene expression profile of a
XX tissue or cell sample exposed to the compound, and comparing the gene
XX expression profile to a database comprising at least part of the data or
XX information given in the specification. The methods are useful for
XX predicting at least one toxic effect of a compound, predicting the
XX progression of a toxic effect of a compound, predicting the renal
XX toxicity of a compound, or identifying toxicity markers in tissues or
XX cells exposed to known renal toxin. The genes are useful as toxicity
XX markers in drug screening and toxicity assays, in monitoring disease or
XX physiological states, or disease progression. This polynucleotide
XX describes a rat DNA sequence relating to the toxic effect database
XX described in the specification. NOTE: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the World Intellectual Property
XX Organization
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 28.3%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.4e-14;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCGAGGAGCATTTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACCC 247
DB 8 CACAAACCGAGGAGCATTTGAACACTGCACAGCGCATCTGCCAGAGAGCTGTGACCC 67

QY 248 CACTTCGCTACTAGCTA 265
DB 68 CACTTCGCTACTACTA 85

RESULT 10
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
XX ADP72689;
XX
XX 26-AUG-2004 (first entry)
XX
XX Renal toxin progression gene marker #1278.
XX
XX
```


KW ds: toxic effect; gene expression profile; kidney tissue;
 KW differential gene expression; toxicity progression; toxicity marker;
 KW drug screening; toxicity assay; kidney pathology; nephritis;
 KW kidney necrosis; glomerular injury; tubular injury;
 KW focal segmental glomerulosclerosis.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2004048598-A2.
 PD 10-JUN-2004.
 XX
 PF 24-NOV-2003; 2003WO-US037556.
 XX
 PR 22-NOV-2002; 2002US-00301856.
 XX
 PA (GENE-) GENE LOGIC INC.
 PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B,
 PI Elashoff M;
 XX
 DR WPI; 2004-460771/43.
 XX
 PT Predicting (the progression of) a toxic effect of a compound, for
 PT monitoring the progression of renal disease states, comprises preparing a
 PT gene expression profile of a kidney tissue or cell sample exposed to the
 PT compound.
 PS Claim 11; SEQ ID NO 1278; 266pp; English.
 XX
 CC The invention relates to a method of predicting (the progression of) a
 CC toxic effect of a compound by preparing a gene expression profile of a
 CC kidney tissue or cell sample exposed to the compound and comparing the
 CC gene expression profile to a database, or detecting the level of gene(s)
 CC expression in a tissue or cell sample exposed to the compound, where
 CC differential gene expression compared to a control indicates a toxic
 CC effect (toxicity progression). The method is useful for predicting (the
 CC progression of) at least one toxic effect of a compound. The genes are
 CC useful as toxicity markers in drug screening and toxicity assays. The
 CC methods are useful for predicting the likelihood that a compound or test
 CC agent will induce various specific kidney pathologies, such as nephritis,
 CC kidney necrosis, glomerular and tubular injury, or focal segmental
 CC glomerulosclerosis. The methods are useful for determining the similarity
 CC of a toxic response to one or more individual compounds and for
 CC predicting or elucidating the potential cellular pathways influenced,
 CC induced or modulated by the compound or test agent. The kit is useful for
 CC predicting or modelling the toxic response of a test compound, for
 CC monitoring the progression of renal disease states, for identifying genes
 CC that show promise as new drug targets and for screening known and newly
 CC designed drugs. This sequence corresponds to a gene marker used in the
 CC method of the invention. (Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 28.3%; Score 76.4; DB 12; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.4e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 188 CACAAACCCAGGAGCATTTGAACTGACACGGCCATCTGCGCAGAGAGCTGTGACAC 247
 DB 8 CACAAACCCAGGAGCATTTGAACTGACACGGCCATCTGCGCAGAGAGCTGTGACAC 67
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 68 CACTTCGGCTACTAGCTA 85
 RESULT 11
 AAH22429
 ID AAH22429 standard; DNA; 5001 BP.
 XX

AC AAH22429;
 XX
 DT 22-AUG-2001 (first entry)
 XX
 DE Rat insulin-like growth factor binding protein nucleotide sequence.
 XX
 KW Identification; toxic; hepatotoxic; differential gene expression; NSAID;
 KW non-steroidal antiinflammatory drug; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO200138579-A2.
 PD 31-MAY-2001.
 XX
 PF 21-NOV-2000; 2000WO-US032049.
 XX
 PR 22-NOV-1999; 99US-0166923P.
 PR 18-FEB-2000; 2000US-0183531P.
 PR 20-NOV-2000; 2000US-00717321.
 XX
 PA (CURA-) CURAGEN CORP.
 PI Gould-Rothberg BE, DiIppio VA, Rameeh TM, Gerwein RW;
 PI WPI; 2001-355948/37.
 DR
 XX
 PT Screening hepatotoxic agent comprises contacting test cell population
 PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
 PT with reference population and identifying difference in expression
 PT levels.
 PS Disclosure; Page 22-24; 76pp; English.
 XX
 CC The present invention describes a method of screening a test agent for
 CC hepatotoxicity. The method comprises: (a) providing a test cell
 CC population comprising a cell capable of expressing one or more nucleic
 CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
 CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
 CC agent; (c) measuring expression of one or more of the nucleic acid
 CC sequences in the test cell population; (d) comparing the expression of
 CC the nucleic acid sequence in the test cell population to the expression
 CC of the nucleic acid sequence in a reference cell population comprising
 CC at least one cell whose exposure status to a hepatotoxic agent is known;
 CC and (e) identifying a difference in expression levels of the RISKMARKER
 CC or INJURYMARKER sequences, if present, in the test cell population and
 CC reference cell population. The method is useful for identifying a
 CC hepatotoxic agent. The present sequence is given in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;
 Query Match 28.3%; Score 76.4; DB 4; Length 5001;
 Best Local Similarity 98.7%; Pred. No. 5.3e-14;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 188 CACAAACCCAGGAGCATTTGAACTGACACGGCCATCTGCGCAGAGAGCTGTGACAC 247
 DB 1 CACAAACCCAGGAGCATTTGAACTGACACGGCCATCTGCGCAGAGAGCTGTGACAC 60
 QY 248 CACTTCGGCTACTAGCTA 265
 DB 61 CACTTCGGCTACTAGCTA 78
 RESULT 12
 ABN95896
 ID ABN95896 standard; DNA; 6128 BP.
 AC ABN95896;
 XX
 DT 13-AUG-2002 (first entry)
 XX

Best Local Similarity 66.5%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACAGCGGATCTTACGACGCTGCCCTGACATCAT 112
Db 473 CACTACCAAAACAACTTATTTTGAACAGCTCCTTACGCGCGCTGCAATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCGCAGCCCTTCAATAAGGCCCTGGGTATGGCAGC 162
Db 533 TAACCTCTCGTGCAAGTGGCGCGCGCTGTGCTTATTAAGGTGGCGCTGTGTCCAGC 592
QY 163 CAGCATGTGCTCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 221
Db 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTCCCGCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCAACCACTTC 253
Db 645 CCACCTCCCGAGAGAGACTGCGCACCGCTCC 676

RESULT 14

AAH57489
ID AAH57489 standard; cDNA; 6134 BP.

AC AAH57489;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human liver cell specific cDNA sequence SEQ ID NO:329.
XX
XX Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
XX
XX Human; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
XX
XX metabolic disease; developmental disease; cytotoxic; immunomodulatory;
XX
XX neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
XX Homo sapiens.
XX
XX OS
XX PN MO200132927-A2.
XX
XX PD 10-MAY-2001.
XX
XX PF 02-NOV-2000; 2000MO-US030396.
XX
XX PR 04-NOV-1999; 99US-0163508P.
XX
XX PA (INCY-) INCYTE GENOMICS INC.
XX
XX PI Sornasee T, Seilhamer J, Watson GA;
XX
XX DR MPI; 2001-291057/30.
XX
XX PT New cell and tissue specific polynucleotides useful for diagnosis,
XX
XX PT prognosis or monitoring of treatments for disorders where the gene is
XX
XX PT associated with a cancer, immunopathology or neuropathology.
XX
XX PS Claim 1; Page 246-248; 327pp; English.

AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
sequences (I). (I) can have cytosolic, immunomodulatory and
neuroprotective activities, and can be used in gene therapy. (I) and
proteins (II) encoded by them are used in high throughput screening
assays to select DNA molecules, RNA molecules, peptide nucleic acids,
mimetics, peptides, proteins, agonists, antagonists, antibodies or their
fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
agents. Expression of (I) in a sample indicates the differentiation of
embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
XX

Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;

Query Match 22.5%; Score 60.8; DB 4; Length 6134;

Best Local Similarity 66.5%; Pred. No. 6.5e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACAGCGGATCTTACGACGCTGCCCTGACATCAT 112
Db 473 CACTACCAAAACAACTTATTTTGAACAGCTCCTTACGCGCGCTGCAATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCGCAGCCCTTCAATAAGGCCCTGGGTATGGCAGC 162
Db 533 TAACCTCTCGTGCAAGTGGCGCGCGCTGTGCTTATTAAGGTGGCGCTGTGTCCAGC 592
QY 163 CAGCATGTGCTCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 221
Db 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTCCCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCAACCACTTC 253
Db 645 CCACCTCCCGAGAGAGACTGCGCACCGCTCC 676

RESULT 15

ABX15371
ID ABX15371 standard; DNA; 51 BP.

AC ABX15371;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat glucose response element (GIRE) DNA.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGRBP-1;
XX
XX insulin-sensitive element; ISR; basal promoter; hyperglycaemia; insulin;
XX
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
XX
XX OS
XX PN US2002107198-A1.
XX
XX PD 08-AUG-2002.
XX
XX PF 10-OCT-2001; 2001US-00972916.
XX
XX PR 11-OCT-2000; 2000US-0239113P.
XX
XX PA (THUL/) THULE P M.
XX
XX PI Thule PM;
XX
XX DR MPI; 2002-674190/72.
XX
XX PT New insulin regulator cassette, useful e.g. for treating diabetes,
XX
XX PT provides specific, glucose-inducible transgenic expression of insulin in
XX
XX PT liver cells.

Claim 7; Page 13; 37pp; English.

The invention relates to an insulin regulator construct comprising a
glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
promoter and an insulin-sensitive element (ISR) of an insulin-like growth
factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
to treat or prevent diabetic complications, to regulate insulin
production, to modulate hyperglycaemia (without severe hypoglycaemia), to
increase fat catabolism and to retard protein catabolism. The construct
provides stimulation of insulin expression by glucose (but not other
carbohydrates) and glucocorticoids and inhibition by glucagon so that a
combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat glucose response element of the invention
 XX

SQ Sequence 51 bp; 6 A; 18 C; 17 G; 10 T; 0 U; 0 Other;

Query Match 18.9%; Score 51; DB 6; Length 51;

Best Local Similarity 100.0%; Pred. No. 1.7e-06;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGGGACTCCCGTGGTTCCTGACTCTGGCCCCCAGTGTA 51
 |||||||
 DB 1 CATGGCGCACGGGGGACTCCCGTGGTTCCTGACTCTGGCCCCCAGTGTA 51

Search completed: September 1, 2005, 00:34:43
 Job time : 288.304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 31, 2005, 12:52:03 ; Search time 340.384 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321

Sequence: 1 taacactggggccagagtc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2002as:*
6: geneseqn2003as:*
7: geneseqn2004as:*
8: geneseqn2005as:*
9: geneseqn2006as:*
10: geneseqn2007as:*
11: geneseqn2008as:*
12: geneseqn2009as:*
13: geneseqn2010as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	321	100.0	321	6	ABX15374 Rat insul
2	321	100.0	372	6	ABX15375 Rat insul
3	219.4	68.3	270	6	ABX15373 Rat insul
4	219.4	68.3	423	6	ABX15376 Rat insul
5	219	68.2	219	6	ABX15372 Rat insul
6	103	32.1	423	6	ABX15376 Rat insul
7	76.4	23.8	1500	6	ABK63701 Rat seque
8	76.4	23.8	1500	10	ADBS8201 Toxicity-
9	76.4	23.8	1500	10	ADBS2710 Primary r
10	76.4	23.8	1500	10	ABT41911 Toxicity
11	76.4	23.8	1500	12	ADP72689 Renal tox
12	76.4	23.8	5001	4	AAH22429 Rat insul
13	63.8	19.9	6128	6	ABN95896 Gene #239
14	63.8	19.9	6128	10	AAH22429 Rat insul
15	63.8	19.9	6134	4	AAH57489 Human liv
16	55.2	17.2	1304	2	AAH57489 Human liv
17	55.2	17.2	13011	2	AAH57489 Human liv
18	55.2	17.2	13011	10	ABT42448 Toxicity
19	55.2	17.2	13011	12	ADP72914 Renal tox
20	51	15.9	51	6	ABX15380 Rat liver

ALIGNMENTS

RESULT 1	ABX15374	standard; DNA; 321 BP.
ID	ABX15374	standard; DNA; 321 BP.
AC	ABX15374;	
XX		
DT	17-APR-2003	(first entry)
DE	Rat insulin regulator construct DNA #2.	
KW	Glucose response element; GIRE; liver pyruvate kinase; Lpk; ds; IGF1P-1;	Abx15371 Rat gluco
KW	insulin-sensitive element; ISE; basal promoter; hyperglycemia; insulin;	Abx15373 Rat insul
KW	insulin-like growth factor binding protein-1; hypoglycemia; glucose;	Abx15379 Rat liver
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;	Ab132997 Human imm
KW	glucagon; euglycemia; diabetes; fasting; ketogenesis; ketoadicosis;	Ab170248 Chemical1
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;	AA661189 Human gen
KW	intracellular hormone receptor; insulin regulator construct; anabolic.	Ad552441 Human cod
XX		
OS	Rattus norvegicus.	Ab970481 Human bon
OS	Synthetic.	Ab190266 Human pol
PN	US2002107198-A1.	Ab132997 Human end
XX		
PD	08-AUG-2002.	Ab132997 Human end
XX		
PF	10-OCT-2001; 2001US-00972916.	Ab132997 Human end
XX		
PR	11-OCT-2000; 2000US-0239113P.	Ab132997 Human end
XX		
PA	(THUL/) THUL P M.	Ab132997 Human end
XX		
PI	Thule PM;	Ab132997 Human end
XX		
DR	WPI; 2002-674190/72.	Ab132997 Human end
XX		
PT	New insulin regulator cassette, useful e.g. for treating diabetes,	Ab132997 Human end
PT	provides specific, glucose-inducible transgenic expression of insulin in	Ab132997 Human end
PT	liver cells.	Ab132997 Human end
PS	Claim 9; Page 14; 37p; English.	Ab132997 Human end
XX		
CC	The invention relates to an insulin regulator construct comprising a	Ab132997 Human end
CC	glucose response element (GIRE) of a liver pyruvate kinase (Lpk) gene	Ab132997 Human end

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 XX

SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 100.0%; Score 321; DB 6; Length 321;

Best Local Similarity 100.0%; Pred. No. 4,4e-89;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCACTGGGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 Db 1 TCACTGGGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 QY 61 GGCAGAGTCCAGGAACCAAGGAGTCCCGGCCCATGTACAGCAAACT 120
 Db 61 GGCAGAGTCCAGGAACCAAGGAGTCCCGGCCCATGTACAGCAAACT 120
 QY 121 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 180
 Db 121 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 180
 QY 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCGCCGAGACA 240
 Db 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCGCCGAGACA 240
 QY 241 CAAACCCAGCAGCATTTGAACACTGACACCGGCATCTGCCAGAGAGTGTGACACCA 300
 Db 241 CAAACCCAGCAGCATTTGAACACTGACACCGGCATCTGCCAGAGAGTGTGACACCA 300
 QY 301 CTTCCGCTACTAGTAGCCGC 321
 Db 301 CTTCCGCTACTAGTAGCCGC 321

RESULT 2

ABX15375
 ID ABX15375 standard; DNA; 372 BP.

AC ABX15375;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #3.

XX Glucose response element; GIRE; liver pyruvate kinase, LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.
 OS Synthetic.

XX US2002107198-A1.

XX 08-AUG-2002.
 PD

XX 10-OCT-2001; 2001US-00972916.
 PF 11-OCT-2000; 2000US-0239113P.
 PR (THULI/) THULE P M.
 PA Thule PM;
 PI WPI; 2002-674190/72.
 DR

XX New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 XX liver cells.
 PS Claim 9; Page 14; 37pp; English.

XX The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 XX

SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 321; DB 6; Length 372;

Best Local Similarity 100.0%; Pred. No. 4,7e-89;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCACTGGGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 60
 Db 52 TCACTGGGGGGCCAGAGTCCAGAACCAAGGAGTCCCGGCCCATGTACACTGGG 111
 QY 61 GGCAGAGTCCAGGAACCAAGGAGTCCCGGCCCATGTACAGCAAACT 120
 Db 112 GGCAGAGTCCAGGAACCAAGGAGTCCCGGCCCATGTACAGCAAACT 171
 QY 121 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 180
 Db 121 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 180
 QY 172 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 231
 Db 172 TATTTGAACAGGGGATCTCTAGCAGCTGCGCCCTGACATTCATTACCCGTGCGAG 231
 QY 181 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCGCCGAGACA 240
 Db 232 CCAGCCCTTCATTAAAGCCCTGGGTATGGCCAGCCAGCATGTCCACTGCGCCGAGACA 291
 QY 241 CAAACCCAGCAGCATTTGAACACTGACACCGGCATCTGCCAGAGAGTGTGACACCA 300
 Db 292 CAAACCCAGCAGCATTTGAACACTGACACCGGCATCTGCCAGAGAGTGTGACACCA 351
 QY 301 CTTCCGCTACTAGTAGCCGC 321
 Db 352 CTTCCGCTACTAGTAGCCGC 372

RESULT 3

ABX15373

ID ABX15373 standard; DNA; 270 BP.

XX ABX15373;
 AC

```
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
KM Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KM intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
XX Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THUL P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 9; Page 13-14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetic during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
XX
XX Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;
XX
XX Query Match 68.3%; Score 219.4; DB 6; Length 270;
XX Best Local Similarity 99.5%; Pred. No. 1.1e-57;
XX Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 101 TGTGACAGCAAAACAACTTATTTGAAACGCGGGATCTTAGACAGCTGCCCTGACANT 160
XX |TGTGACAGCAAAACAACTTATTTGAAACGCGGGATCTTAGACAGCTGCCCTGACANT 160
DB 50 TATCAACAACAAACAACTTATTTGAAACGCGGGATCTTAGACAGCTGCCCTGACANT 109
XX
XX 161 CATTACCCGCTGCTGCGGAGCCAGCCCTTCATTAAGCCCTCGGTAGTGGCAACGACGANT 220
XX |CATTACCCGCTGCTGCGGAGCCAGCCCTTCATTAAGCCCTCGGTAGTGGCAACGACGANT 220
DB 110 CATTACCCGCTGCTGCGGAGCCAGCCCTTCATTAAGCCCTCGGTAGTGGCAACGACGANT 169
XX
XX 221 GTTCACATGCGCGCGGAGACAAACCCAGAGCATTTGAACACTGCACAGGCAATCTGC 280
XX |GTTCACATGCGCGCGGAGACAAACCCAGAGCATTTGAACACTGCACAGGCAATCTGC 280
DB 170 GTTCACATGCGCGCGGAGACAAACCCAGAGCATTTGAACACTGCACAGGCAATCTGC 229
```

```
QY 281 CCAGAGAGCTGTGACACCACTTCGCGTACTAGCTAGCGCG 321
XX |CCAGAGAGCTGTGACACCACTTCGCGTACTAGCTAGCGCG 321
DB 230 CCAGAGAGCTGTGACACCACTTCGCGTACTAGCTAGCGCG 270
XX
XX RESULT 4
XX ID ABX15376 standard; DNA; 423 BP.
XX
XX ABX15376;
XX
XX 17-APR-2003 (first entry)
XX
XX Rat insulin regulator construct DNA #4.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
XX
XX Synthetic.
XX
XX US2002107198-A1.
XX
XX 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THUL P M.
XX
XX Thule PM;
XX
XX WPI; 2002-674190/72.
XX
XX New insulin regulator cassette, useful e.g. for treating diabetes,
XX provides specific, glucose-inducible transgenic expression of insulin in
XX liver cells.
XX
XX Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetic during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a rat insulin regulator construct of the invention
XX
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
XX
XX Query Match 68.3%; Score 219.4; DB 6; Length 423;
XX Best Local Similarity 99.5%; Pred. No. 1.3e-57;
XX Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	101	NGTCAAGCAAAACAACTATTATTTTGAACAGCGGGATCCTACAGCTGCGCTTGACAAT	160
Db	203	TATCAACAAGCAAAACAACTATTATTTTGAACAACGGGAGTCCCTGACACT	262
QY	161	CATTAAACCGTCTGCTCGAGCCAGCCCTTCATTAAGGCCCTGGGTATATGGCCAGCAGCATG	220
Db	263	CATTAAACCGTCTGCTCGAGCCAGCCCTTCATTAAGGCCCTGGGTATATGGCCAGCAGCATG	322
QY	221	GTCACCTGCGCGCCGAGACACAACCCAGCGAGCATTTGAACACTGACACAGGCCATCTGC	280
Db	323	GTCCACTGCGCCCGCAGACACAACCCAGCGAGCATTTGAACACTGACACAGGCCATCTGC	382
QY	261	CCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTAGCCGC	321
Db	363	CCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTAGCCGC	423
RESULT 5			
ABX15372	ABX15372 standard; DNA; 219 BP.		
XX	ABX15372;		
AC	17-APR-2003	(first entry)	
XX	Rat insulin-sensitive element (ISE) DNA.		
DE	Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;		
KW	insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;		
KW	insulin-like growth factor binding protein-1; hypoglycaemia; glucose;		
KW	fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;		
KW	glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;		
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;		
KW	intracellular hormone receptor; insulin regulator construct; anabolic.		
XX	Rattus norvegicus.		
XX	US2002107198-A1.		
PN	08-AUG-2002.		
XX	10-OCT-2001; 2001US-00972916.		
PD	11-OCT-2000; 2000US-0239113P.		
XX	(THUL/) THULE P M.		
PA	Thule PM;		
XX	New insulin regulator cassette, useful e.g. for treating diabetes,		
DR	provides specific, glucose-inducible transgenic expression of insulin in		
XX	liver cells.		
PT	Claim 8; Page 13; 37pp; English.		
PT	The invention relates to an insulin regulator construct comprising a		
CC	glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene		
CC	promoter and an insulin-sensitive element (ISE) of an insulin-like growth		
CC	factor binding protein-1 (IGFBP-1) basal promoter. The construct is used		
CC	to treat or prevent diabetic complications, to regulate insulin		
CC	production, to modulate hyperglycaemia (without severe hypoglycaemia), to		
CC	increase fat catabolism and to retard protein catabolism. The construct		
CC	provides stimulation of insulin expression by glucose (but not other		
CC	carbohydrates) and glucocorticoids and inhibition by glucagon so that a		
CC	combination of these effects can maintain nearly euglycaemic conditions		
CC	in diabetics during short-term fasting, large carbohydrate loads or when		
CC	fed ad libitum and prevent pathological ketogenesis and ketoacidosis,		
CC	thus inhibiting the long-term complications of diabetes. The properties		
CC	of the construct are essentially specific for hepatocytes and well-		
CC	differentiated hepatoma lines and insulin expression in these cells may		
CC	have effects additional to those provided by secreted insulin, e.g.		

CC		inhibition of cellular protein degradation, and inhibition, or
CC		stimulation of other intracellular hormone receptors. This sequence
CC		represents a rat insulin-sensitive element of the invention
XX		
SQ		Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;
<hr/>		
	Query Match	68.2%; Score 219; DB 6; Length 219;
	Best Local Similarity	100.0%; Pred. No. 1,4e-57;
	Matches 219; Conservative	0; Mismatches 0; Indels 0; Gaps 0
<hr/>		
QY		103 TCACAGAAAAAACAATTATTTTAAACAGGGGATCTTAGCACCCTGCCATCATCA 162
Db		1 TCACAGAAAACAAACTTATTTTAAACAGGGGATCTTAGCACCCTGCCATCATCA 60
<hr/>		
OY		163 TTAACTCCGTGTCGCAGCACGCCCCCTTCATAAGGCTTGATATGCGACGACATGCT 222
Db		61 TTAACTCCGTGTCGCAGCACGCCCCCTTCATAAGGCTTGATATGCGACGACATGCT 120
<hr/>		
OY		223 CCACATGCCCCCGGAGAACCAAAACCCAGAGACATTTGAACACTGCACACGGGCATGTGCC 282
Db		121 CCACATGCCCCCGGAGAACCAAAACCCAGAGACATTTGAACACTGCACACGGGCATGTGCC 180
<hr/>		
OY		283 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTAGCGCG 321
Db		181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTAGCGCG 219
<hr/>		
RESULT 6		
ABX15376/C		
ID	ABX15376	standard; DNA; 423 BP.
XX		
AC	ABX15376;	
XX		
DT	17-APR-2003	(first entry)
XX		
DE	Rat insulin regulator construct DNA #4.	
XX		
KW	Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;	
KW	insulin-sensitive element; ISB; basal promoter; hyperglycaemia; insulin;	
KW	insulin-like growth factor binding protein-1; hypoglycaemia; glucose;	
KW	fat caraboliem; protein caraboliem; carbohydrate; glucocorticoid; rat;	
KW	glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;	
KW	hepatocyte; hepatoma; cellular protein degradation; antidiabetic;	
KW	intracellular hormone receptor; insulin regulator construct; anabolic.	
XX		
OS	Rattus norvegicus.	
OS	Synthetic.	
XX		
PN	US2002107198-A1.	
XX		
PD	08-AUG-2002.	
XX		
PF	10-OCT-2001; 2001US-00972916.	
XX		
PR	11-OCT-2000; 2000US-0239113P.	
XX		
PA	(THUL/) THULB P M.	
PI	Thule PM;	
XX		
DR	WPI; 2002-674190/72.	
XX		
PT	New insulin regulator cassette, useful e.g. for treating diabetes,	
PT	provides specific, glucose-inducible transgenic expression of insulin in	
XX	liver cells.	
XX		
PS	Claim 9; Page 14; 37pp; English.	
CC		The invention relates to an insulin regulator construct comprising a
CC		glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC		promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC		factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC		to treat or prevent diabetic complications, to regulate insulin

CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
 Query Match 32.1%; Score 103; DB 6; Length 423;
 Best Local Similarity 100.0%; Pred. No. 1.2e-21;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TACACTGGGGGCGCAGATCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 60
 DB 204 TACACTGGGGGCGCAGATCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 145
 QY 61 GGCCAGATCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 103
 DB 144 GGCCAGATCCAGAACCCAGGAGTCCCGTCCGCCATGTACTGAGG 102
 RESULT 7
 ABK63701
 ID ABK63701 standard; cDNA; 1500 BP.
 AC ABK63701;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
 KW Rat; sg; hepatotoxin; expressed sequence tag; EST; drug screening;
 XX differential expression; centrilobular necrosis; steatosis.
 OS Rattus norvegicus.
 XX
 PN WO200210453-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 30-JUL-2001; 2001WO-US023872.
 XX
 PR 31-JUL-2000; 2000US-0222040P.
 XX
 PR 02-NOV-2000; 2000US-0244880P.
 XX
 PR 11-MAY-2001; 2001US-0290029P.
 XX
 PR 15-MAY-2001; 2001US-0290645P.
 XX
 PR 22-MAY-2001; 2001US-0292336P.
 XX
 PR 06-JUN-2001; 2001US-0295798P.
 XX
 PR 13-JUN-2001; 2001US-0297457P.
 XX
 PR 19-JUN-2001; 2001US-0298884P.
 XX
 PR 09-JUL-2001; 2001US-0303459P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter WM, Johnson KR, Caetle AL, Elashoff MR;
 XX
 DR WPI; 2002-241625/29.
 XX
 PT Predicting toxic effects of compounds or the progression of these toxic
 XX effects by determining the changes in gene expression in tissues or cells
 PT exposed to the toxin and comparing these to gene expression in unexposed
 XX tissues or cells.
 PS Claim 1; SEQ ID NO 1608; 239pp; English.

XX The invention relates to methods for predicting toxic effects of
 CC compounds or the progression of these toxic effects by determining the
 CC global changes in gene expression in tissues or cells exposed to the
 CC toxin and comparing these to gene expression in unexposed tissues or
 CC cells. Also included are methods of predicting at least one toxic effect
 CC of a compound or progression of a toxic effect, preferably the
 CC hepatotoxicity of a compound, comprising detecting the level of
 CC expression in a tissue or cell sample exposed to the compound of two or
 CC more genes listed in the specification, where differential expression of
 CC the genes is indicative of at least one toxic effect or progression. The
 CC method can also be used to identify an agent which modulates the toxic
 CC response and predict cellular pathways that a compound modulates in a
 CC cell. The methods utilize a set of at least two probes (on a solid
 CC support in kit form), where each of the probes comprises a sequence that
 CC specifically hybridises to a gene listed in the specification, a computer
 CC system comprising a database containing information identifying the
 CC expression level in a tissue or cell sample exposed to a hepatotoxin of a
 CC set of genes comprising at least two genes listed in the specification,
 CC and a user interface to view the information used to present information
 CC identifying the expression level in a tissue or cell of at least one gene
 CC listed in the specification. The method is useful for elucidating global
 CC changes in gene expression and for identifying toxicity markers in
 CC tissues or cell exposed to a known toxin. The genes may be used as
 CC toxicity markers in drug screening and toxicity assays. The genes and
 CC gene expression information may be used as diagnostic markers for the
 CC prediction or identification of the physiological state of tissue or cell
 CC sample that has been exposed to a compound or agent. Hepatotoxicity is
 CC characterised by centrilobular necrosis and steatosis. The present
 CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene
 CC which is differentially expressed in response to a hepatotoxic agent
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 6; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTGAACACTGCACAGGCGCATCTGCCAGAGCTGTGACCAC 298
 DB 8 CACAAACCCAGGAGCATTGAACACTGCACAGGCGCATCTGCCAGAGCTGTGACCAC 67
 QY 299 CACTTCGCGTACTACTA 316
 DB 68 CACTTCGCGTACTACTA 85
 RESULT 8
 ADB58201
 ID ADB58201 standard; DNA; 1500 BP.
 AC ADB58201;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Toxicity-related gene, SEQ ID 3227.
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 XX drug screening; toxicity assay; de.
 OS Unidentified.
 XX
 PN WO2003064624-A2.
 XX
 PD 07-AUG-2003.
 XX
 PF 31-JAN-2003; 2003WO-US003194.
 XX
 PR 31-JAN-2002; 2002US-00060087.
 XX
 PR 15-MAR-2002; 2002US-0364045P.
 XX
 PR 15-MAR-2002; 2002US-0364055P.
 XX
 PR 30-DEC-2002; 2002US-0436643P.
 XX

PA (GENE-) GENE LOGIC INC.
 XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 PI WPI, 2003-689530/65.
 XX
 PT Predicting a toxic effect of a compound, useful in identifying toxicity
 markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 the compound.
 XX
 PS Claim 1; SEQ ID NO 3227, 1156bp; English.
 XX
 CC The present invention relates to a method for predicting a toxic effect
 of a compound. The method comprises preparing a gene expression profile
 of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 298
 DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
 QY 299 CACTTCGCTACTATCTA 316
 DB 68 CACTTCGCTACTATCTA 85
 Db
 RESULT 9
 ID ADB52710 standard; DNA; 1500 BP.
 XX
 AC ADB52710;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
 XX
 KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
 KW toxicity marker; toxicity progression; drug screening;
 KW primary rat hepatocyte toxicity modelling; gene; ds.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2003065993-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 04-FEB-2003; 2003WO-US003482.
 XX
 PR 04-FEB-2002; 2002US-0353171P.
 PR 13-MAR-2002; 2002US-036534P.
 PR 08-APR-2002; 2002US-0370248P.
 PR 10-APR-2002; 2002US-0371134P.
 PR 10-APR-2002; 2002US-0371135P.

PR 10-APR-2002; 2002US-0371150P.
 PR 11-APR-2002; 2002US-0371413P.
 PR 19-APR-2002; 2002US-0373601P.
 PR 19-APR-2002; 2002US-0373602P.
 PR 22-APR-2002; 2002US-0374139P.
 PR 08-MAY-2002; 2002US-0378370P.
 PR 09-MAY-2002; 2002US-0378652P.
 PR 09-MAY-2002; 2002US-0378653P.
 PR 09-MAY-2002; 2002US-0378655P.
 PR 09-JUL-2002; 2002US-0394230P.
 PR 09-JUL-2002; 2002US-0394253P.
 PR 04-SEP-2002; 2002US-0407688P.
 PR 28-JAN-2003; 2003US-0442900P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
 PI Elashoff M;
 XX
 DR WPI, 2003-731472/69.
 XX
 PT Determining if a compound induces a toxic effect on a tissue or cell, for
 PT identifying hepatotoxic compounds, comprises comparing a gene expression
 PT profile of a tissue or cell sample to a database of tox mean and non-tox
 PT mean values.
 XX
 PS Claim 44; SEQ ID NO 3252; 874bp; English.
 XX
 CC The present invention describes a method for determining whether a
 CC compound induces a toxic effect on a tissue or cell. The method comprises
 CC preparing a gene expression profile of a tissue or cell sample exposed to
 CC the compound, and comparing the gene expression profile to a database
 CC comprising data or information on the tox mean and non-tox mean value.
 CC The method is useful for predicting or identifying at least one toxic
 CC effect, particularly hepatotoxicity, of a test or unknown compound. The
 CC genes listed in the specification are useful as diagnostic or toxicity
 CC markers for the prediction or identification of the physiological state
 CC of tissue or cell sample that has been exposed to a compound, or to
 CC identify or predict the toxic effects of a compound or an agent. These
 CC may also be used as markers for monitoring toxicity progression or for
 CC drug screening. The present sequence represents a primary rat hepatocyte
 CC toxicity modelling related gene sequence from the present invention.
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 Query Match 23.8%; Score 76.4; DB 10; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3.2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 239 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 298
 DB 8 CACAAACCCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
 QY 299 CACTTCGCTACTATCTA 316
 DB 68 CACTTCGCTACTATCTA 85
 Db
 RESULT 10
 ID ABT41911 standard; DNA; 1500 BP.
 XX
 AC ABT41911;
 XX
 DT 26-JUN-2003 (first entry)
 XX
 DE Toxicity modelling related rat gene SEQ ID No 1613.
 XX
 KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
 KW database; drug screening; toxicity assay; rat; ds.
 XX
 OS Rattus norvegicus.
 XX

PN WO200295000-A2.
XX
XX 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
XX 22-MAY-2001; 2001US-0292335P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 01-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-036144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
PA
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M,
XX
XX WPI; 2003-148464/14.
XX
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modeling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX
XX Example 4; Page; 446pp; English.
XX
XX
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
XX
XX
XX Query Match 23.8%; Score 76.4; DB 10; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3,2e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY 239 CACAAACCAGCGAGCATTTGACACACATGTGACACGCGCATTCGCCAGAGAGCTGTGACAC 298
DB 8 CACAAACCAGCGAGCATTTGACACACATGTGACACGCGCATTCGCCAGAGAGCTGTGACAC 67
QY 299 CACTTCGGCTACTACTA 316
DB 68 CACTTCGGCTACTACTA 85

RESULT 11
 ADP72689
 ID ADP72689 standard; DNA; 1500 BP.
 XX
 XX ADP72689;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Renal toxin progression gene marker #1278.
 XX
 KW ds; toxic effect; gene expression profile; kidney tissue;
 KW differential gene expression; toxicity progression; toxicity marker;
 KW drug screening; toxicity assay; kidney pathology; nephritis;
 KW kidney necrosis; glomerular injury; tubular injury;
 KW focal segmental glomerulosclerosis.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2004048598-A2.
 XX
 PD 10-JUN-2004.
 XX
 PF 24-NOV-2003; 2003WO-US037556.
 XX
 PR 22-NOV-2002; 2002US-00301856.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
 PI Elashoff M;
 XX
 DR WPI; 2004-460771/43.
 XX
 PT Predicting (the progression of) a toxic effect of a compound, for
 PT monitoring the progression of renal disease states, comprises preparing a
 PT gene expression profile of a kidney tissue or cell sample exposed to the
 PT compound.
 XX
 PS Claim 11; SEQ ID NO 1278; 266pp; English.
 XX
 XX The invention relates to a method of predicting (the progression of) a
 XX toxic effect of a compound by preparing a gene expression profile of a
 XX kidney tissue or cell sample exposed to the compound and comparing the
 XX gene expression profile to a database, or detecting the level of gene(s)
 XX differential in a tissue or cell sample exposed to the compound, where
 XX differential gene expression compared to a control indicates a toxic
 XX effect (toxicity progression). The method is useful for predicting (the
 XX progression of) at least one toxic effect of a compound. The genes are
 XX useful as toxicity markers in drug screening and toxicity assays. The
 XX methods are useful for predicting the likelihood that a compound or test
 XX agent will induce various specific kidney pathologies, such as nephritis,
 XX kidney necrosis, glomerular and tubular injury, or focal segmental
 XX glomerulosclerosis. The methods are useful for determining the similarity
 XX of a toxic response to one or more individual compounds and for
 XX predicting or elucidating the potential cellular pathways influenced,
 XX induced or modulated by the compound or test agent. The kit is useful for
 XX predicting or modelling the toxic response of a test compound, for
 XX monitoring the progression of renal disease states, for identifying genes
 XX that show promise as new drug targets and for screening known and newly
 XX designed drugs. This sequence corresponds to a gene marker used in the
 XX method of the invention. (Note: The sequence data for this patent did not
 XX form part of the invention. (Note: The sequence data for this patent did not
 XX format directly from WIPO at ftp.wipo.int/pub/published_pcc_sequences)).
 XX
 SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
 XX
 Query Match 23.8%; Score 76.4; DB 12; Length 1500;
 Best Local Similarity 98.7%; Pred. No. 3,2e-13;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Db 8 CACAAACCCAGGACATTGAACTGCAACAGCCATCTGCCAGAGCTGTGACAC 67
OY 239 CACTTCGGCTACTAGCTA 316
|||
68 CACTTCGGCTACTAGCTA 85
|||
RESULT 12
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
AC AAH22429;
XX
DT 22-AUG-2001 (first entry)
XX
DE Rat insulin-like growth factor binding protein nucleotide sequence.
KM Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KM non-steroidal antiinflammatory drug; ds.
XX
OS Rattus norvegicus.
XX
PN WO200138579-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032049.
XX
PR 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gould-Rothberg BE, Dipippo VA, Ramseh TM, Gerwein RW;
XX
XX WPI; 2001-355948/37.
XX
PT Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INTURMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
PS Disclosure; Page 22-24; 76pp; English.
XX
CC The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INTURMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in an reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INTURMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

Query Match 23.8%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 4.7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 239 CACAAACCCAGGACATTGAACTGCAACAGCCATCTGCCAGAGCTGTGACAC 298
|||
Db 1 CACAAACCCAGGACATTGAACTGCAACAGCCATCTGCCAGAGCTGTGACAC 60
|||
OY 239 CACTTCGGCTACTAGCTA 316
|||

Db 61 CACTTCGGCTACTAGCTA 78
RESULT 13
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
AC ABN95896;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2394 used to diagnose liver cancer.
XX
KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KM metastatic liver tumour; cytostatic; expression profile; disease state;
KM disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
XX WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
XX
PS Claim 1; SEQ ID NO 2394; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cyrostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;

Query Match 19.9%; Score 63.8; DB 6; Length 6128;
Best Local Similarity 63.2%; Pred. No. 3.9e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

OY 69 TCCAGAAACCCAGGAGTGGCCCGTCGCGCCATGTCAAGCAAAACTTATTTTGA 128
|||
Db 438 TCCATCCCAACGACGGGTTTGGCTAGGGCTTGGGTCACATGCAAAACTTATTTTGA 497
|||
OY 129 ACAGCGGAGATTGACAGCGTGCCTTGACAATCTTAACCC-----GTGCTGGCG 178
|||
Db 498 ACACTCAGCTCTTAGCGTGGCGGCTCTCCATCTTAACCTCTGTGCAAGTGCGCGG 557
|||
OY 179 AGCCAGCCCTTCAAAAGCCCTGGTATGGCAAGCCGAGCATGTGCCCTGCGCGCGAGA 238
|||
Db 558 CTTGTGCTTTATTAAGGTGGCGCTGTGTCCAGGAGCATGGCCACCGCATTC--- 613
|||

QY 239 CACAAACCCAGCAGATTGTAACACTGC-ACACGCCATCTGCCCAGAGAGCTGTGACCA 297
 DB 614 ----CATCCAGCAGACATCTGCGCCGCCGCCGCCGCCACCTCCAGAGAGCACTGGCCA 669
 QY 298 CCACTTC 304
 DB 670 CCGCTCC 676

RESULT 14
 ABV75371
 ID ABV75371 standard; DNA; 6128 BP.
 AC ABV75371;
 XX
 DT 07-MAR-2003 (first entry)
 DE Human IGFBP-1 gene sequence.
 XX
 KW Insulin-like growth factor binding protein; IGFBP; cytosstatic; liver;
 XX cancer; human; IGFBP-1; gene; db.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 762..5200
 FT /tag= a
 FT /product= "IGFBP-1"
 FT /note= "Insulin-like growth factor binding protein;
 FT contains introns"
 FT exon 762..1110
 FT /tag= b
 FT /number= 1
 FT /tag= c
 FT /number= 1
 FT /tag= d
 FT /number= 2
 FT /tag= e
 FT /number= 2
 FT /tag= f
 FT /number= 3
 FT /tag= g
 FT /number= 3
 FT /tag= h
 FT /number= 4
 XX
 XX WO200290580-A1.
 XX PN 14-NOV-2002.
 XX PD 03-MAY-2002; 2002WO-AU000558.
 XX PF 03-MAY-2001; 2001US-0288441P.
 XX PR 03-MAY-2001; 2001US-0288441P.
 XX
 PA (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
 PA (SICE-) SINGAPORE GEN HOSPITAL PTE LTD.
 PA (ARGA/) ARGENT V P.
 XX
 XX Huynh TH, Chow PKH, Soo KC;
 XX
 DR MPI; 2003-103522/09.
 DR P-PDB; ABB82757.
 XX
 PT Detecting the presence or diagnosing the risk of a liver cancer in a
 PT patient comprises detecting aberrant expression of a gene encoding an
 PT insulin-like growth factor binding protein.
 XX

PS Example; Page 104-108; 142pp; English.
 XX
 CC The invention relates to detecting the presence or diagnosing the risk of
 CC a liver cancer in a patient. The method involves detecting in a
 CC biological sample obtained from the patient aberrant expression of a gene
 CC encoding an insulin-like growth factor binding protein (IGFBP). The
 CC method is useful for detecting the presence or diagnosing the risk of a
 CC liver cancer or for screening agents in a patient. The agent is useful
 CC for the manufacture of a medicament for treating and/or preventing liver
 CC cancer. The present sequence represents a human IGFBP-1 polypeptide
 CC encoding genomic DNA (GenBank Accession No. M74587)
 XX
 SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
 Query Match 19.9%; Score 63.8; DB 10; Length 6128;
 Best Local Similarity 63.2%; Pred. No. 3.9e-09;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
 QY 69 TCACAGAACACAGGAGTGCCTCCGCCGCCCATGTCACAGCAAAACAACTATTATTCGA 128
 DB 438 TCCTCCACACAGGAGTTCCTGAGGCTTGAGGCTGCTGACAGCAAAACAACTATTATTCGA 497
 QY 129 ACACGGGATCCTAGCAGCTGCCCTGACAAATCATTAACCC-----GTGCTGCCG 178
 DB 498 ACATCAGCTCCTACGCTGCGCGCTGCCATCATTAACCTCTGTCGCAAGTGGCGCG 557
 QY 179 AGCCAGCCCTTCATTAAGGCCCTGGGTATGCGCCAGCAGCATGTCACCTGCCGCCGAGA 238
 DB 558 CCGTGGCCCTTATTAAGGTGCGCGCTGTGTCACAGCAGCATGCGCACCGCCATCC---- 613
 QY 239 CACAAACCCAGCAGATTGTAACACTGC-ACACGCCATCTGCCAGAGAGCTGTGACCA 297
 DB 614 ----CATCCAGCAGACATCTGCGCCGCCGCCGCCGCCACCTCCAGAGAGCACTGGCCA 669
 QY 298 CCACTTC 304
 DB 670 CCGCTCC 676

RESULT 15
 AAH57489
 ID AAH57489 standard; cDNA; 6134 BP.
 XX
 AC AAH57489;
 XX
 DT 10-SEP-2001 (first entry)
 DE Human liver cell specific cDNA sequence SEQ ID NO:329.
 XX
 KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
 KW liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytosstatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
 XX
 OS Homo sapiens.
 XX PN WO200132927-A2.
 XX PD 10-MAY-2001.
 XX PF 02-NOV-2000; 2000WO-US030396.
 XX PR 04-NOV-1999; 99US-0163508P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Sorrasse T, Selthamer J, Watson GA;
 XX
 DR MPI; 2001-291057/30.
 XX
 PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is
 PT associated with a cancer, immunopathology or neuropathology.
 XX

XX Claim 1; Page 246-248; 327pp; English.

CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
 CC sequences (I). (I) can have cytosolic, immunomodulatory and
 CC neuroprotective activities, and can be used in gene therapy. (I) and
 CC proteins (II) encoded by then are used in high throughput screening
 CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
 CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
 CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
 CC agents. Expression of (I) in a sample indicates the differentiation of
 CC embryonic stem cells into a tissue selected from brain, heart, kidney,
 CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
 CC to produce an expression profile that defines a metabolic or
 CC developmental process, treatment, condition, disease or disorder. The
 CC gene profile can be used for diagnosis, prognosis or monitoring of
 CC treatments and for investigating a predisposition to a disorder where the
 CC gene is associated with a cancer, immunopathology or neuropathology
 XX

SQ Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;

Query Match 19.9%; Score 63.8; DB 4; Length 6134;
 Best Local Similarity 63.2%; Pred. No. 3.9e-09;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Oy	69	TCCAGGAAACGAGGAGTCCCGGTGCGCCATGTCAAGAGCAAAACAACTTATTTGA	128
Db	438	TCTTCCACACGACGGTTTGCCTAGGSCCTTGGGTGACATGCAAAACAACTTATTTGA	497
Oy	129	ACACGGGATCTTACGACGCTGCTGACATCATTAACCC-----GTGCTGCCG	178
Db	498	ACACTCAGTCTTACGCTGCGGCTGCAATCATTAACCTCTGTGCAAGTGGCGCG	557
Oy	179	AGCCAGCCCTTATTAAGCCCTGGTATGCGCAGCCAGATGTGCATCTCCCGCGAGA	238
Db	558	CCTGTCCCTTATTAAGTGGCGCTGTGTCCAGCGAGCATGCGCACCGCATCC----	613
Oy	239	CACAAACCGAGGAGATGAACACTGC-ACAGGCCATCTGCCAGAGAGCTGTGACCA	297
Db	614	----CATTCAGGAGCATCTGCCCGCGCCCGCCACCTCTCCAGAGAGCATGGCCA	669
Oy	298	CCACTTC	304
Db	670	CCGCTCC	676

Search completed: September 1, 2005, 00:34:45
 Job time : 342.384 secs

CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene

ABX15376
ID ABX15376 standard; DNA; 423 BP.
XX
AC ABX15376;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #4.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
XX US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
XX (THUL/) THULE P M.
XX
PA Thule PM;
XX
PI WPI; 2002-674190/72.
XX
DR New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
XX
Query Match 60.1%; Score 223.6; DB 6; Length 423;
Best Local Similarity 76.2%; Pred. No. 8,8e-58;
Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

QY 122 CAGGAACCAAGGAGTGCCTTGCCTGAGCCATG-----TCACAAACA 162
DB 154 CATGGGCGCAGCGGGGACATCCCGTGTCTCTGAGACTTGAGCCCAAGTGTATCACAAACA 213
QY 163 AAACAAACTTATTTTGAACACGCGGATCTTACAGACGTCGCCCTGACAAATCATTTAACCGT 222
DB 214 AAACAACTTATTTTGAACACGCGGATCTTACAGACGTCGCCCTGACAAATCATTTAACCGT 273
QY 223 GCTGCCAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGCATGTCTACATGCC 282
DB 274 GCTGCCAGCCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCAGCATGTCTACATGCC 333
QY 283 GCCGAGACCAAAACCCAGCAGCATTTGAACACTGCACACGCGCATCTGCCAGAGACTG 342
DB 334 GCCGAGACCAAAACCCAGCAGCATTTGAACACTGCACACGCGCATCTGCCAGAGACTG 393
QY 343 TGACCAACCACTTCGCTACTAGCTAGCCGC 372
DB 394 TGACCAACCACTTCGCTACTAGCTAGCCGC 423

RESULT 4
ABX15373
ID ABX15373 standard; DNA; 270 BP.
XX
AC ABX15373;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
XX Rattus norvegicus.
OS Synthetic.
OS
XX US2002107198-A1.
XX
PN 08-AUG-2002.
XX
PD 10-OCT-2001; 2001US-00972916.
XX
PF 11-OCT-2000; 2000US-0239113P.
XX
PR (THUL/) THULE P M.
XX
PA Thule PM;
XX
PI WPI; 2002-674190/72.
XX
DR New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 13-14; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketocidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match 59.0%; Score 219.4; DB 6; Length 270;
Best Local Similarity 99.5%; Pred. No. 1.4e-56;

Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 152 TGTCAAGCAAAACAACTTATTTGAAACAGGGAGATCTTACACGCTGCTTACAT 211
DB 50 TATCAAGCAAAACAACTTATTTGAAACAGGGAGATCTTACACGCTGCTTACAT 109

QY 212 CATTAACCGGCTGCGGAGACAGCCCTTCATTAAGCCCTGGGTATGGCCAGACGATG 271

DB 110 CATTAACCGGCTGCGGAGACAGCCCTTCATTAAGCCCTGGGTATGGCCAGACGATG 169

QY 272 GTTCACCTGCGCCGAGACAAACCCAGAGCATTTGAACATGCAACGCGCATCTGC 331

DB 170 GTTCACCTGCGCCGAGACAAACCCAGAGCATTTGAACATGCAACGCGCATCTGC 229

QY 332 CCAAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCG 372

DB 230 CCAAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCG 270

RESULT 5

ABX15372
ID ABX15372 standard; DNA; 219 BP.

AC ABX15372;

DT 17-APR-2003 (first entry)

DE Rat insulin-sensitive element (ISE) DNA.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketocidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

PA (THUL/) THULE P M.

PI Thule PM;

DR WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

PS Claim 8; Page 13; 37pp; English.

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketocidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention
XX

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 58.9%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.8e-56;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TCACAGCAAAACAACTTATTTGAAACAGGGAGATCTTACACGCTGCTTACATCA 213

DB 1 TCACAGCAAAACAACTTATTTGAAACAGGGAGATCTTACACGCTGCTTACATCA 60

QY 214 TTAACCCGCTGCTGCGGAGACAGCCCTTCATTAAGCCCTGGGTATGGCCAGCAGCATG 273

DB 61 TTAACCCGCTGCTGCGGAGACAGCCCTTCATTAAGCCCTGGGTATGGCCAGCAGCATG 120

QY 274 CCACTGCCCGCCGAGACAAACCCAGAGCATTTGAACATGCAACGCGCATCTGCC 333

DB 121 CCACTGCCCGCCGAGACAAACCCAGAGCATTTGAACATGCAACGCGCATCTGCC 180

QY 334 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCG 372

DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGCG 219

RESULT 6

ABX15376/C
ID ABX15376 standard; DNA; 423 BP.

AC ABX15376;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #4.

KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketocidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

PA (THUL/) THULE P M.

```

XX Thule PW;
PI MPI; 2002-674190/72.
DR MPI; 2002-674190/72.
XX New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX Claim 9; Page 14; 37pp; English.
XX The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;
SQ
Query Match 41.4%; Score 154; DB 6; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.2e-36;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TACACTGGGGGGCCAGAGTCCAGAAACCGGAGTCCCGGCCCATGTACACTGGG 60
DB 204 TACACTGGGGGGCCAGAGTCCAGAAACCGGAGTCCCGGCCCATGTACACTGGG 145
QY 61 GCGCAGAGTCCAGAACCAAGGAGTGCCTCGGCCCATGTACACTGGGGCCAGAGT 120
DB 144 GCGCAGAGTCCAGAACCAAGGAGTGCCTCGGCCCATGTACACTGGGGCCAGAGT 85
QY 121 CCAGAACCCAGCGGAGTGCCTCGGCCCATGT 154
DB 84 CCAGAACCCAGCGGAGTGCCTCGGCCCATGT 51
RESULT 7
ABK63701
ID ABK63701 standard; cDNA; 1500 BP.
XX
AC ABK63701;
XX
DT 18-JUN-2002 (first entry)
XX
DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
XX Rat; as; hepatotoxin; expressed sequence tag; EST; drug screening;
XX differential expression; centrilobular necrosis; steatosis.
XX
OS Rattus norvegicus.
XX
PN WO200210453-A2.
XX
PD 07-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-US023872.
XX
PR 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.

```

```

PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX MPI; 2002-241625/29.
XX
DR Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
XX method can also be used to identify an agent which modulates the toxic
XX response and predict cellular pathways that a compound modulates in a
XX cell. The method utilizes a set of at least two probes (on a solid
XX support in kit form), where each of the probes comprises a sequence that
XX specifically hybridises to a gene listed in the specification, a computer
XX system comprising a database containing information identifying the
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX set of genes comprising at least two genes listed in the specification,
XX and a user interface to view the information used to present information,
XX identifying the expression level in a tissue or cell of at least one gene
XX listed in the specification. The method is useful for elucidating global
XX changes in gene expression and for identifying toxicity markers in
XX tissues or cell exposed to a known toxin. The genes may be used as
XX toxicity markers in drug screening and toxicity assays. The genes and
XX gene expression information may be used as diagnostic markers for the
XX prediction or identification of the physiological state of tissue or cell
XX sample that has been exposed to a compound or agent. Hepatotoxicity is
XX characterised by centrilobular necrosis and steatosis. The present
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX which is differentially expressed in response to a hepatotoxic agent
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
Query Match 20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAACCCAGCGAGTGAACACTGCACAGGCGCATCTCCAGAGAGCTGTACAC 349
DB 8 CACAACCCAGCGAGTGAACACTGCACAGGCGCATCTCCAGAGAGCTGTACAC 67
QY 350 CACTTCGCGTACTAGCTA 367
DB 68 CACTTCGCGTACTATCTA 85
RESULT 8
ADB58201
ID ADB58201 standard; DNA; 1500 BP.
XX
AC ADB58201;
XX

```

DT 04-DEC-2003 (first entry)
XX Toxicity-related gene, SEQ ID 3227.
DE
XX
XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KM drug screening; toxicity assay; ds.
XX
XX Unidentified.
OS
XX WO2003064624-A2.
XX
XX
XX 07-AUG-2003.
PD
XX
XX 31-JAN-2003; 2003WO-US003194.
PF
XX
XX 31-JAN-2002; 2002US-00060087.
PR 15-MAR-2002; 2002US-0364045P.
PR 15-MAR-2002; 2002US-0364055P.
PR 30-DEC-2002; 2002US-0436643P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M,
PI WPI; 2003-689530/65.
XX
XX
XX Predicting a toxic effect of a compound, useful in identifying toxicity
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
the compound.
XX
XX Claim 1; SEQ ID NO 3227; 1156bp; English.
XX
XX The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound, and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
CC least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAAACCCGAGCATTTGAACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 349
DB 8 CACAAACCCGAGCATTTGAACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 350 CACTTCGCTACTAGCTA 367
DB 68 CACTTCGCTACTAGCTA 85
RESULT 9
ID ADB52710 standard; DNA; 1500 BP.
XX ADB52710;
AC
XX
DT 04-DEC-2003 (first entry)

XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX Rattus norvegicus.
OS
XX
XX WO2003065993-A2.
XX
XX
XX 14-AUG-2003.
PD
XX
XX 04-FEB-2003; 2003WO-US003482.
PF
XX
XX 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M,
PI Elashoff M;
XX WPI; 2003-731472/69.
XX
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX
XX Claim 44; SEQ ID NO 3252; 874bp; English.
XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
SQ
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 290 CACAAACCCGAGCATTTGAACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 349
DB 8 CACAAACCCGAGCATTTGAACTGACACGCGCATCTGCCAGAGAGCTGTGACAC 67

Qy 350 CACTTCGGCTACTAGCTA 367
|||||
Db 68 CACTTCGGCTACTATCTA 85

```
RESULT 10
ABT41911
ID ABT41911 standard; DNA; 1500 BP.
XX
AC ABT41911;
XX
DT 26-JUN-2003 (first entry)
XX
DE Toxicity modelling related rat gene SEQ ID No 1613.
XX
KM Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KW database; drug screening; toxicity assay; rat; ds.
XX
OS Rattus norvegicus.
XX
PN WO200295000-A2.
XX
PD 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
PR 22-MAY-2001; 2001US-0299233P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-SEP-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0366134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI
DR WPI; 2003-148464/14.
XX
XX Predicting at least one toxic effect of a compound, useful for toxicity
PT modelling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
XX Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
CC effect of a compound. The method comprises a gene expression profile of a
CC tissue or cell sample exposed to the compound, and comparing the gene
CC expression profile to a database comprising at least part of the data or
CC information given in the specification. The methods are useful for
CC predicting at least one toxic effect of a compound, predicting the renal
CC progression of a toxic effect of a compound, predicting the renal
CC toxicity of a compound, or identifying toxicity markers in tissues or
CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
```

```
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
Query Match 20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
Qy 230 CACAACCAGGAGCATTTGAACACTGCACGCGCATCTGCGCAGAGCTGTACAC 349
|||||
Db 8 CACAACCAGGAGCATTTGAACACTGCACGCGCATCTGCGCAGAGCTGTACAC 67
|||||
Qy 350 CACTTCGGCTACTAGCTA 367
|||||
Db 68 CACTTCGGCTACTATCTA 85
|||||
XX
XX RESULT 11
XX ADP72689
XX ID ADP72689 standard; DNA; 1500 BP.
XX
XX AC ADP72689;
XX
XX DT 26-AUG-2004 (first entry)
XX
XX DE Renal toxin progression gene marker #1278.
XX
XX ds; toxic effect; gene expression profile; kidney tissue;
XX differential gene expression; toxicity progression; toxicity marker;
XX drug screening; toxicity assay; kidney pathology; nephritis;
XX kidney necrosis; glomerular injury; tubular injury;
XX focal segmental glomerulosclerosis.
XX
XX OS Rattus norvegicus.
XX
XX PN WO2004048598-A2.
XX
XX PD 10-JUN-2004.
XX
XX PF 24-NOV-2003; 2003WO-US037556.
XX
XX PR 22-NOV-2002; 2002US-00301856.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
XX PI Elashoff M;
XX
XX DR WPI; 2004-460771/43.
XX
XX Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
XX Claim 11; SEQ ID NO 1278; 266pp; English.
XX
XX The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The
CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
```

CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 20.5%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7e-13; 1; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCAGGAGCATTTGAACACTGCACAGGCGCATCTGCCAGAGAGCTGTGACAC 349
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACAGGCGCATCTGCCAGAGAGCTGTGACAC 67
QY 350 CACTTCGGCTACTAGCTA 367
DB 68 CACTTCGGCTACTAGCTA 85

RESULT 12

AAH22429
ID AAH22429 standard; DNA; 5001 BP.

XX AAH22429;

DT 22-AUG-2001 (first entry)

DE Rat insulin-like growth factor binding protein nucleotide sequence.

KM Identification; toxic; hepatotoxic; differential gene expression; NSAIID;

KW non-steroidal antiinflammatory drug; ds.

XX Rattus norvegicus.

XX WO200138579-A2.

PD 31-MAY-2001.

PF 21-NOV-2000; 2000WO-US032049.

PR 22-NOV-1999; 99US-0166923P.

PR 18-FEB-2000; 2000US-0183531P.

PR 20-NOV-2000; 2000US-00717321.

XX (CURA-) CURAGEN CORP.

XX Gould-Rothberg BE, Dipippo VA, Rameeh TM, Gerwein RW;

XX WPI; 2001-355948/37.

PT Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.

XX Disclosure; Page 22-24; 76pp; English.

XX The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test
CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression

CC of the nucleic acid sequence in an reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention
XX

SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;

Query Match 20.5%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 1e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCAGGAGCATTTGAACACTGCACAGGCGCATCTGCCAGAGAGCTGTGACAC 349
DB 1 CACAAACCCAGGAGCATTTGAACACTGCACAGGCGCATCTGCCAGAGAGCTGTGACAC 60
QY 350 CACTTCGGCTACTAGCTA 367
DB 61 CACTTCGGCTACTAGCTA 78

RESULT 13

ABN95896
ID ABN95896 standard; DNA; 6128 BP.

XX ABN95896;

DT 13-AUG-2002 (first entry)

DE Gene #2394 used to diagnose liver cancer.

KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

KW metastatic liver tumour; cytostatic; expression profile; disease state;

KW disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

PD 11-APR-2002.

PF 02-OCT-2001; 2001WO-US030589.

PR 02-OCT-2000; 2000US-0237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvarez C, Peres-Da-Silva S, Vockley JG;

XX WPI; 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.

XX Claim 1; SEQ ID NO 2394; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
 SQ
 Query Match 17.2%; Score 63.8; DB 6; Length 6128;
 Best Local Similarity 63.2%; Pred. No. 7.3e-09;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
 QY 120 TCCAGAACCCAGGAGTCCCGGCGCCCATGTGACAGAAACAACTTATTTGA 179
 DB 438 TCCCTCCACAGCGGTTGCGTAGAGGCTTGAGTGCACAGAAACAACTTATTTGA 497
 QY 180 ACACGGGGATCTAGACGCTGCCCTGACATCATTAACCC-----GTGCTGCCG 229
 DB 498 ACACCTAGCTCTAGGCTGCGCGCTGCAATCATTAACCTCTGAGTGCAGAGTGGCGG 557
 QY 230 AGCCAGCCCTTATTAAGCCCTGAGTATGCGCAGCAGCATGATGCTGCGCGGAGA 289
 DB 558 CCTGTGCCCTTATTAAGTGGCGCGCTGTGTCCAGAGAGATGCGCACCGCATTC---- 613
 QY 290 CACAAACCCAGCGAGCATTTGAACACTGC-ACACGGCCATCTGCCAGAGAGCTGTGACCA 348
 DB 614 ----CATCCAGCGAGCATCTGCCGCGCGCCGCCACCTCCAGAGAGCATGTGCCA 669
 QY 349 CCACCTTC 355
 DB 670 CCGCTCC 676
 RESULT 14
 ABV75371
 ID ABV75371 standard; DNA; 6128 BP.
 XX
 AC ABV75371;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE Human IGFBP-1 gene sequence.
 XX
 KW Insulin-like growth factor binding protein; IGFBP; cytosolic; liver;
 KW cancer; human; IGFBP-1; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS /tag= a
 FT /product= "IGFBP-1"
 FT /note= "Insulin-like growth factor binding protein;
 FT contains introns"
 FT exon 762..1110
 FT /tag= b
 FT /number= 1
 FT intron 1111..2656
 FT /tag= c
 FT /number= 1
 FT exon 2657..2826
 FT /tag= d
 FT /number= 2
 FT intron 2827..4040
 FT /tag= e
 FT /number= 2
 FT exon 4041..4169
 FT /tag= f
 FT /number= 3
 FT intron 4170..5068
 FT /tag= g
 FT /number= 3
 FT exon 5069..5197
 FT /tag= h
 FT /number= 4
 PN WO200290580-A1.

XX
 PD 14-NOV-2002.
 XX
 XX 03-MAY-2002; 2002WO-AU000558.
 PF
 PR 03-MAY-2001; 2001US-0288441P.
 XX
 PA (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
 PA (SICE-) SINGAPORE GEN HOSPITAL PTE LTD.
 PA (ARGA/) ARGAE V P.
 XX
 PI Huynh TH, Chow PKH, Soo KC;
 XX
 DR WPI, 2003-103522/09.
 DR P-PSDB; ABB82757.
 XX
 PT Detecting the presence or diagnosing the risk of a liver cancer in a
 PT patient comprises detecting aberrant expression of a gene encoding an
 PT insulin-like growth factor binding protein.
 XX
 PS Example; Page 104-108; 142pp; English.
 CC The invention relates to detecting the presence or diagnosing the risk of
 CC a liver cancer in a patient. The method involves detecting in a
 CC biological sample obtained from the patient aberrant expression of a gene
 CC encoding an insulin-like growth factor binding protein (IGFBP). The
 CC method is useful for detecting the presence or diagnosing the risk of a
 CC liver cancer or for screening agents in a patient. The agent is useful
 CC for the manufacture of a medicament for treating and/or preventing liver
 CC cancer. The present sequence represents a human IGFBP-1 polypeptide
 CC encoding genomic DNA (Genbank Accession No. M74587)
 XX
 SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
 XX
 Query Match 17.2%; Score 63.8; DB 10; Length 6128;
 Best Local Similarity 63.2%; Pred. No. 7.3e-09;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
 QY 120 TCCAGAACCCAGGAGTCCCGGCGCCCATGTGACAGAAACAACTTATTTGA 179
 DB 438 TCCCTCCACAGCGGTTGCGTAGAGGCTTGAGTGCACAGAAACAACTTATTTGA 497
 QY 180 ACACGGGGATCTAGACGCTGCCCTGACATCATTAACCC-----GTGCTGCCG 229
 DB 498 ACACCTAGCTCTAGGCTGCGCGCTGCAATCATTAACCTCTGAGTGCAGAGTGGCGG 557
 QY 230 AGCCAGCCCTTATTAAGCCCTGAGTATGCGCAGCAGCATGATGCTGCGCGGAGA 289
 DB 558 CCTGTGCCCTTATTAAGTGGCGCGCTGTGTCCAGAGAGATGCGCACCGCATTC---- 613
 QY 290 CACAAACCCAGCGAGCATTTGAACACTGC-ACACGGCCATCTGCCAGAGAGCTGTGACCA 348
 DB 614 ----CATCCAGCGAGCATCTGCCGCGCGCCGCCACCTCCAGAGAGCATGTGCCA 669
 QY 349 CCACCTTC 355
 DB 670 CCGCTCC 676
 RESULT 15
 AAH57489
 ID AAH57489 standard; CDNA; 6134 BP.
 XX
 AC AAH57489;
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human liver cell specific CDNA sequence SEQ ID NO:329.
 XX
 KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
 KW liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.

```
XX Homo sapiens.
OS
XX WO200132927-A2.
PN
XX 10-MAY-2001.
PD
XX 02-NOV-2000; 2000WO-US030396.
PF
XX 04-NOV-1999; 99US-0163508P.
PR
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Sornasse T, Seilhamer JJ, Watson GA;
PI
XX WPI; 2001-291057/30.
DR
XX
XX
XX New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology.
XX
XX Claim 1; Page 246-248; 327pp; English.
PS
XX AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytosolic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC agents. Expression of (I) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
XX
SQ Sequence 6134 BP; 1505 A; 1498 C; 1507 G; 1624 T; 0 U; 0 Other;
Query Match 17.2%; Score 63.8; DB 4; Length 6134;
Best Local Similarity 63.2%; Pred. No. 7.3e-09;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;
QY 120 TCCAGAGAACCAAGGAGTCCCGCCGATGTCACAGCAAACTATTGGA 179
DB |||||
438 TCCTCCACACAGCGTTGCGTAGGACCTGGGTGCACAGCAAACTATTGGA 497
QY 180 ACACGGGGATCCTAGCAGCTGCCCTGACATCATTAACCC-----GTGCTGCCG 229
DB |||||
498 ACACCTCAGCTCTTACGTCGCGCGCTGCCAATCATTAACCTCTGTGCAAGTGGCGCG 557
QY 230 AGCCAGCCCTTCATTAAGCGCTGGGTATGCGCAGCAGCATGTCCACTGCCCGCGAGA 289
DB |||||
558 CCTGTGCCCTTATTAAGTGCAGCTGTGTCCAGCAGCATGTGCGCACCGCATCC---- 613
QY 290 CACAACCCAGGAGCATTTGAACATGCG-ACACGGCCATGTGCCAGAGAGCTGTGACCA 348
DB |||||
614 ---CATCAGAGGAGCATCTGCGCGCGCGCGCCACCCCTCCAGAGAGCATGTGCCA 669
QY 349 CCACTTC 355
DB |||||
670 CCGCTCC 676
```

Search completed: September 1, 2005, 00:34:47
Job time : 396.464 secs

CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 SQ Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 423; DB 6; Length 423;

Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGC 60
 DB 1 CATGGCGCAGCGGCGACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGC 60
 QY 61 ACGGGGCACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGCGAC 120
 DB 61 ACGGGGCACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGCGAC 120
 QY 121 TCCCGGTGTTCTGTGACTCTGCGCCCCAGTGTATAGTGGCGCGACTCCCGTGT 180
 DB 121 TCCCGGTGTTCTGTGACTCTGCGCCCCAGTGTATAGTGGCGCGACTCCCGTGT 180
 QY 181 TCTGTGACTCTGCGCCCCAGTGTATAGTGGCGCGAACTTATTTTGAACGCGGAT 240
 DB 181 TCTGTGACTCTGCGCCCCAGTGTATAGTGGCGCGAACTTATTTTGAACGCGGAT 240
 QY 241 CCTAGACGCTGCGCCCCAGTGTATAGTGGCGCGAACTTATTTTGAACGCGGAT 300
 DB 241 CCTAGACGCTGCGCCCCAGTGTATAGTGGCGCGAACTTATTTTGAACGCGGAT 300
 QY 301 CTGGGTATGGCGCGAGCATGTGTCACGCGCGCGGAGACCAACCCAGGAGCATTTG 360
 DB 301 CTGGGTATGGCGCGAGCATGTGTCACGCGCGCGGAGACCAACCCAGGAGCATTTG 360
 QY 361 AACACTGCACAGCGGCATCTGCGCGAGAGAGTGTGACCACTTCCGCTACTAGTAGC 420
 DB 361 AACACTGCACAGCGGCATCTGCGCGAGAGAGTGTGACCACTTCCGCTACTAGTAGC 420
 QY 421 CGC 423
 DB 421 CGC 423

RESULT 2

ABX15373

ID ABX15373 standard; DNA; 270 BP.

AC ABX15373;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #1.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
 XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;

KW intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 XX Synthetic.
 XX US2002107198-A1.
 XX 08-AUG-2002.
 XX 10-OCT-2001; 2001US-00972916.
 XX 11-OCT-2000; 2000US-0239113P.
 XX (THULE/) THULE P M.
 XX Thule PM;
 XX WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 13-14; 37pp; English.

The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX
 SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;

Query Match Best Local Similarity 100.0%; Score 270; DB 6; Length 270;

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 CATGGCGCAGCGGCGACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGC 213
 DB 1 CATGGCGCAGCGGCGACTCCCGTGGTCTGTGACTCTGCCCCCAGTGTATAGTGGCGC 60
 QY 214 AAACAACTTATTTTGAACAGCGGATCTGTGACGCTGCGCTTGAACATCTTAACCGT 273
 DB 61 AAACAACTTATTTTGAACAGCGGATCTGTGACGCTGCGCTTGAACATCTTAACCGT 120
 QY 274 GCTGCGAGCGAGCCCTTCAATAGGCCCTGGGTATGGCCAGCAGCATGTGTCATGCGC 333
 DB 121 GCTGCGAGCGAGCCCTTCAATAGGCCCTGGGTATGGCCAGCAGCATGTGTCATGCGC 180
 QY 334 GCCGAGACAAACCCAGCGGATTTGAACACTGACACGCGCATCTTCCAGAGAGCTG 393
 DB 181 GCCGAGACAAACCCAGCGGATTTGAACACTGACACGCGCATCTTCCAGAGAGCTG 240
 QY 394 TGACCAACCTTCGCTACTAGTAGCGC 423
 DB 241 TGACCAACCTTCGCTACTAGTAGCGC 270

RESULT 3

ABX15375
 ID ABX15375 standard; DNA; 372 BP.
 AC ABX15375;
 XX
 XX
 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #3.
 XX
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KM Insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KM Insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX
 XX US2002107198-A1.
 PN
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 XX Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetes during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 XX
 XX Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
 XX
 Query Match 52.9%; Score 223.6; DB 6; Length 372;
 Best Local Similarity 76.2%; Pred. No. 4.3e-59;
 Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;
 QY 34 ACTCTGGCCCGCCAGGTATACATGCGGCACTCCCGTTCCTGAGCTCTGGCC 93
 DB 2 AACTCTGGGCGCAAGTCTCCAGGAACACACGGAGTGCCTCCGCGCCCACTGACCTGGGG 61
 QY 94 CCCAGTGTATCATGGCGCACTCCCGTGTCTCTGAGCTCTGGCCCGCCAGTGTGA 153
 DB 62 GCCAGAGTCTCAGGAACACACGGAGTGCCTCCGCGCCCACTGTAACACTGGGCGCCAGAGTC 121

QY 154 CATGGCGCAGCGGCACTCCCGTGTCTCTGAGCTCTGAGCCCGCCAGTGTATCAGAGCA 213
 DB 122 CAGGAACCACGGAGTGCCTCCGTGCGCCCATG-----TCACAGAGA 162
 QY 214 AAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCGCTGACATCATTAACCGGT 273
 DB 163 AAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCGCTGACATCATTAACCGGT 222
 QY 274 GCTGCCAGCAGCCCTTCATAGGCGCTGGTATGCGCAGCAGCATGCTCACTGCC 333
 DB 223 GCTGCCAGCAGCCCTTCATAGGCGCTGGTATGCGCAGCAGCATGCTCACTGCC 282
 QY 334 GCCGAGACAAACCCAGGAGCATTTGAACATGCAACGCGCATCTGCCGAGAGCTG 393
 DB 283 GCCGAGACAAACCCAGGAGCATTTGAACATGCAACGCGCATCTGCCGAGAGCTG 342
 QY 394 TGACCAACCACTTCGCTACTAGCTAGCGGC 423
 DB 343 TGACCAACCACTTCGCTACTAGCTAGCGGC 372
 RESULT 4
 ID ABX15374 standard; DNA; 321 BP.
 AC ABX15374;
 XX
 XX
 17-APR-2003 (first entry)
 DE Rat insulin regulator construct DNA #2.
 XX
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KM Insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KM Insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 XX
 XX US2002107198-A1.
 PN
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 XX Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions

CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX

SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;

Query Match 51.9%; Score 219.4; DB 6; Length 321;
Best Local Similarity 99.5%; Pred. No. 8.2e-58;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 TATCAGACCAAAACAACTTATTTGAAACAGGGGATCTTGACAGCTGCTGCAAT 262
DB 101 TGTCAACAGCAAAACAACTTATTTGAAACAGGGGATCTTGACAGCTGCTGCAAT 160
QY 263 CATTACCCGCTGCTGCGGAGCCGCTTCATTAAGGCGCTGGGTAATGCGCAGACGATG 322
DB 161 CATTACCCGCTGCTGCGGAGCCGCTTCATTAAGGCGCTGGGTAATGCGCAGACGATG 220
QY 323 GTCCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 382
DB 221 GTCCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 280
QY 383 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCGCG 423
DB 281 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCGCG 321

RESULT 5

ABX15372
ID ABX15372 standard; DNA; 219 BP.

AC ABX15372;

DT 17-APR-2003 (first entry)

DE Rat insulin-sensitive element (ISE) DNA.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; fat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THULE P M.

PA Thule PM;

XX WPI; 2002-674190/72.

PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.

PS Claim 8; Page 13; 37BP; English.

XX

CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin-sensitive element of the invention
XX

SQ Sequence 219 BP; 57 A; 77 C; 48 G; 37 T; 0 U; 0 Other;

Query Match 51.8%; Score 219; DB 6; Length 219;
Best Local Similarity 100.0%; Pred. No. 9.5e-58;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 TCACAGCAAAACAACTTATTTGAAACAGGGGATCTTGACAGCTGCTGCAATCA 264
DB 1 TCACAGCAAAACAACTTATTTGAAACAGGGGATCTTGACAGCTGCTGCAATCA 60
QY 265 TTAACCCGCTGCTGCGGAGCCGCTTCATTAAGGCGCTGGGTAATGCGCAGACGATGT 324
DB 61 TTAACCCGCTGCTGCGGAGCCGCTTCATTAAGGCGCTGGGTAATGCGCAGACGATGT 120
QY 325 CCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGCC 384
DB 121 CCACTGCGCCGCGAGACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGCC 180
QY 385 AGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCGCG 423
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGCTAGCGCG 219

RESULT 6

ABX15375/C
ID ABX15375 standard; DNA; 372 BP.

AC ABX15375;

DT 17-APR-2003 (first entry)

DE Rat insulin regulator construct DNA #3.

XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.

OS Rattus norvegicus.

PN US2002107198-A1.

PD 08-AUG-2002.

PF 10-OCT-2001; 2001US-00972916.

PR 11-OCT-2000; 2000US-0239113P.

XX (THUL/) THULE P M.

PA

XX Thule PM;
 PI
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetes during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;
 Query Match 36.4%; Score 154; DB 6; Length 372;
 Best Local Similarity 100.0%; Pred. No. 1.6e-37;
 Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 ACATGAGGCGACGAGGCGACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 110
 DB 154 ACATGAGGCGACGAGGCGACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 95
 QY 111 CACGGGGGCACTCCGCTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCGGCGCA 170
 DB 94 CACGGGGGCACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCGGCGCA 35
 QY 171 CTCCTGGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCGGCGCA 204
 DB 34 CTCCTGGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCGGCGCA 1
 RESULT 7
 ABX15374/c
 ID ABX15374 standard; DNA; 321 BP.
 XX
 AC ABX15374;
 XX
 DT 17-APR-2003 (first entry)
 XX
 DE Rat insulin regulator construct DNA #2.
 XX
 XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
 KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 OS
 XX US2002107198-A1.
 XX

PD 08-AUG-2002.
 XX
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetes during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 CC
 SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
 Query Match 24.3%; Score 103; DB 6; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 ACATGAGGCGACGAGGCGACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 110
 DB 103 ACATGAGGCGACGAGGCGACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 44
 QY 111 CACGGGGGCACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 153
 DB 43 CACGGGGGCACTCCCGTGTCTCTGACTGTGAGCCCGCACTGATCATGAGCG 1
 RESULT 8
 ABK63701
 ID ABK63701 standard; cDNA; 1500 BP.
 XX
 AC ABK63701;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Rat sequence differentially expressed in response to a hepatotoxin #1608.
 XX
 KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
 KW differential expression; centrilobular necrosis; steatosis.
 XX
 OS Rattus norvegicus.
 OS
 XX WO200210453-A2.
 XX
 PD 07-FEB-2002.
 PD
 PF 30-JUL-2001; 2001WO-US023872.
 PF

```
XX 31-JUL-2000; 2000US-0222040P.
PR 02-NOV-2000; 2000US-0244880P.
PR 11-MAY-2001; 2001US-0290029P.
PR 15-MAY-2001; 2001US-0290645P.
PR 22-MAY-2001; 2001US-0292336P.
PR 06-JUN-2001; 2001US-0295798P.
PR 13-JUN-2001; 2001US-0297457P.
PR 19-JUN-2001; 2001US-0298884P.
PR 09-JUL-2001; 2001US-0303459P.
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;
XX
XX WPI; 2002-241625/29.
XX
XX Predicting toxic effects of compounds or the progression of these toxic
XX effects by determining the changes in gene expression in tissues or cells
XX exposed to the toxin and comparing these to gene expression in unexposed
XX tissues or cells.
XX
XX Claim 1; SEQ ID NO 1608; 239pp; English.
XX
XX The invention relates to methods for predicting toxic effects of
XX compounds or the progression of these toxic effects by determining the
XX global changes in gene expression in tissues or cells exposed to the
XX toxin and comparing these to gene expression in unexposed tissues or
XX cells. Also included are methods of predicting at least one toxic effect
XX of a compound or progression of a toxic effect, preferably the
XX hepatotoxicity of a compound, comprising detecting the level of
XX expression in a tissue or cell sample exposed to the compound of two or
XX more genes listed in the specification, where differential expression of
XX the genes is indicative of at least one toxic effect or progression. The
XX method can also be used to identify an agent which modulates the toxic
XX response and predict cellular pathways that a compound modulates in a
XX cell. The methods utilize a set of at least two probes (on a solid
XX support in kit form), where each of the probes comprises a sequence that
XX specifically hybridises to a gene listed in the specification, a computer
XX system comprising a database containing information identifying the
XX expression level in a tissue or cell sample exposed to a hepatotoxin of a
XX set of genes comprising at least two genes listed in the specification,
XX and a user interface to view the information used to present information
XX identifying the expression level in a tissue or cell of at least one gene
XX listed in the specification. The method is useful for elucidating global
XX changes in gene expression and for identifying toxicity markers in
XX tissues or cell exposed to a known toxin. The genes may be used as
XX toxicity markers in drug screening and toxicity assays. The genes and
XX gene expression information may be used as diagnostic markers for the
XX prediction or identification of the physiological state of tissue or cell
XX sample that has been exposed to a compound or agent. Hepatotoxicity is
XX characterised by centrilobular necrosis and steatosis. The present
XX sequence is an expressed sequence tag (EST) or cDNA derived from a gene
XX which is differentially expressed in response to a hepatotoxic agent
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 6; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 341 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 400
DB 8 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
QY 401 CACTTCGCTACTAGCTA 418
DB 68 CACTTCGCTACTAGCTA 85
RESULT 9
ID ADB58201 standard; DNA; 1500 BP.
```

```
XX ADB58201;
XX
XX 04-DEC-2003 (first entry)
XX
XX Toxicity-related gene, SEQ ID 3227.
XX
XX Toxic; toxin; gene expression profile; hepatotoxicity; liver;
XX drug screening; toxicity assay; ds.
XX
XX Unidentified.
XX
XX WO2003064624-A2.
XX
XX 07-AUG-2003.
XX
XX 31-JAN-2003; 2003WO-US003194.
XX
XX 31-JAN-2002; 2002US-00060087.
XX 15-MAR-2002; 2002US-0364045P.
XX 15-MAR-2002; 2002US-0364055P.
XX 30-DEC-2002; 2002US-0436643P.
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX
XX WPI; 2003-689530/65.
XX
XX Predicting a toxic effect of a compound, useful in identifying toxicity
XX markers in liver tissues or cells for drug screening and toxicity assays,
XX comprises preparing gene expression profile of tissue or cells exposed to
XX the compound.
XX
XX Claim 1; SEQ ID NO 3227; 1156pp; English.
XX
XX The present invention relates to a method for predicting a toxic effect
XX of a compound. The method comprises preparing a gene expression profile
XX of a tissue or cell sample exposed to the compound, and comparing the
XX gene expression profile to a database comprising SEQ ID 1-4925, where
XX differential expression of the gene indicates at least one toxic effect.
XX The method is useful for predicting at least one toxic effect of a
XX compound, predicting hepatotoxicity or the progression of a toxic effect
XX of a compound, identifying an agent that modulates the onset or
XX progression of a toxic response, predicting the cellular pathways that a
XX compound modulates in a cell, and identifying an agent that modulates at
XX least one activity of a protein. The method and compositions of the
XX present invention using a database of genes having liver toxin-induced
XX differential expression, are useful in identifying toxicity markers in
XX liver tissues or cells for drug screening and toxicity assays. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;
XX
XX Query Match 18.1%; Score 76.4; DB 10; Length 1500;
XX Best Local Similarity 98.7%; Pred. No. 3.1e-13;
XX Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 341 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 400
DB 8 CACAAACCCAGGACATTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
QY 401 CACTTCGCTACTAGCTA 418
DB 68 CACTTCGCTACTAGCTA 85
RESULT 10
ID ADB52710 standard; DNA; 1500 BP.
```

AC ADB52710;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3252.
XX
KM toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KM toxicity marker; toxicity progression; drug screening;
KM primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
PN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378655P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/59.
XX
DR Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
PS Claim 44; SEQ ID NO 3252; 874pp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX gene listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX
SO Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

Query Match 18.1%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.1e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

341 CACAAACCCAGCAGCATTTGACACCTGCACACGGCCATCTGCCAGAGACTGTACAC 400

DB 8 CACAAACCCAGCAGCATTTGACACCTGCACACGGCCATCTGCCAGAGACTGTACAC 67
QY 401 CACTTCGCTACTACTA 418
DB 68 CACTTCGCTACTACTA 85
RESULT 11
ABR1911
ID ABR1911 standard; DNA; 1500 BP.
XX
XX ABR1911;
XX
XX 26-JUN-2003 (first entry)
XX
DT
XX
DE Toxicity modelling related rat gene SEQ ID No 1613.
XX
XX
KM Toxic effect; gene expression profile; renal toxicity; toxicity marker;
KM database; drug screening; toxicity assay; rat; ds.
XX
OS Rattus norvegicus.
XX
PN WO200295000-A2.
XX
PD 28-NOV-2002.
XX
PF 22-MAY-2002; 2002WO-US016173.
XX
PR 22-MAY-2001; 2001US-0292335P.
PR 13-JUN-2001; 2001US-0297523P.
PR 19-JUN-2001; 2001US-0298925P.
PR 10-JUL-2001; 2001US-0303807P.
PR 10-JUL-2001; 2001US-0303808P.
PR 10-JUL-2001; 2001US-0303810P.
PR 28-AUG-2001; 2001US-0315047P.
PR 27-SEP-2001; 2001US-0324928P.
PR 22-OCT-2001; 2001US-0330462P.
PR 01-NOV-2001; 2001US-0330867P.
PR 21-NOV-2001; 2001US-0331805P.
PR 06-DEC-2001; 2001US-0336144P.
PR 19-DEC-2001; 2001US-0340873P.
PR 21-FEB-2002; 2002US-0357842P.
PR 21-FEB-2002; 2002US-0357843P.
PR 21-FEB-2002; 2002US-0357844P.
PR 15-MAR-2002; 2002US-0364134P.
PR 08-APR-2002; 2002US-0370144P.
PR 08-APR-2002; 2002US-0370206P.
PR 08-APR-2002; 2002US-0370247P.
PR 17-APR-2002; 2002US-0372794P.
PR 21-APR-2002; 2002US-0371679P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
PI WPI; 2003-148464/14.
XX
XX
DR WPI; 2003-148464/14.
XX
PT Predicting at least one toxic effect of a compound, useful for toxicity
PT modelling, comprises preparing a gene expression profile of a tissue or
PT cell sample exposed to the compound, and comparing the gene expression
PT profile to a database.
XX
PS Example 4; Page; 446pp; English.
XX
XX The invention relates to a novel method of predicting at least one toxic
XX effect of a compound. The method comprises a gene expression profile of a
XX tissue or cell sample exposed to the compound, and comparing the gene
XX expression profile to a database comprising at least part of the data or
XX information given in the specification. The methods are useful for
XX predicting at least one toxic effect of a compound, predicting the
XX progression of a toxic effect of a compound, predicting the renal
XX toxicity of a compound, or identifying toxicity markers in tissues or

CC cells exposed to known renal toxin. The genes are useful as toxicity
CC markers in drug screening and toxicity assays, in monitoring disease or
CC physiological states, or disease progression. This polynucleotide
CC represents a rat DNA sequence relating to the toxic effect database
CC described in the specification. NOTE: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the World Intellectual Property
CC Organization

XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

QY Query Match 18.1%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.1e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 341 CACAAACCCGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 400
8 CACAAACCCGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 401 CACTTCGGCTACTAGCTA 418
|||||
68 CACTTCGGCTACTATCTA 85

Db 68 CACTTCGGCTACTATCTA 85

RESULT 12
ADP72689
ID ADP72689 standard; DNA; 1500 BP.
XX
AC ADP72689;
XX
DT 26-AUG-2004 (first entry)
XX
DE Renal toxin progression gene marker #1278.
XX
KM ds; toxic effect; gene expression profile; kidney tissue;
KM differential gene expression; toxicity progression; toxicity marker;
KM drug screening; toxicity assay; kidney pathology; nephritis;
KM kidney necrosis; glomerular injury; tubular injury;
KM focal segmental glomerulosclerosis.
XX
OS Rattus norvegicus.
XX
PN WO2004048598-A2.
XX
PD 10-JUN-2004.
XX
PF 24-NOV-2003; 2003WO-US037556.
XX
PR 22-NOV-2002; 2002US-00301856.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
PI Elashoff M;
XX
XX WPI; 2004-460771/43.
DR
XX
XX Predicting (the progression of) a toxic effect of a compound, for
PT monitoring the progression of renal disease states, comprises preparing a
PT gene expression profile of a kidney tissue or cell sample exposed to the
PT compound.
XX
XX Claim 11; SEQ ID NO 1278; 266pp; English.

CC The invention relates to a method of predicting (the progression of) a
CC toxic effect of a compound by preparing a gene expression profile of a
CC kidney tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database, or detecting the level of gene(s)
CC expression in a tissue or cell sample exposed to the compound, where
CC differential gene expression compared to a control indicates a toxic
CC effect (toxicity progression). The method is useful for predicting (the
CC progression of) at least one toxic effect of a compound. The genes are
CC useful as toxicity markers in drug screening and toxicity assays. The

CC methods are useful for predicting the likelihood that a compound or test
CC agent will induce various specific kidney pathologies, such as nephritis,
CC kidney necrosis, glomerular and tubular injury, or focal segmental
CC glomerulosclerosis. The methods are useful for determining the similarity
CC of a toxic response to one or more individual compounds and for
CC predicting or elucidating the potential cellular pathways influenced,
CC induced or modulated by the compound or test agent. The kit is useful for
CC predicting or modelling the toxic response of a test compound, for
CC monitoring the progression of renal disease states, for identifying genes
CC that show promise as new drug targets and for screening known and newly
CC designed drugs. This sequence corresponds to a gene marker used in the
CC method of the invention. (Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).

XX
SQ Sequence 1500 BP; 336 A; 408 C; 380 G; 376 T; 0 U; 0 Other;

QY Query Match 18.1%; Score 76.4; DB 12; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.1e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 341 CACAAACCCGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 400
8 CACAAACCCGAGCATTTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 401 CACTTCGGCTACTAGCTA 418
|||||
68 CACTTCGGCTACTATCTA 85

Db 68 CACTTCGGCTACTATCTA 85

RESULT 13
AAH22429
ID AAH22429 standard; DNA; 5001 BP.
XX
XX
AC AAH22429;
XX
DT 22-AUG-2001 (first entry)
XX
DE Rat insulin-like growth factor binding protein nucleotide sequence.
XX
KM Identification; toxic; hepatotoxic; differential gene expression; NSAID;
KM non-steroidal antiinflammatory drug; ds.
XX
OS Rattus norvegicus.
XX
PN WO200138579-A2.
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US032049.
XX
PR 22-NOV-1999; 99US-0166923P.
PR 18-FEB-2000; 2000US-0183531P.
PR 20-NOV-2000; 2000US-00717321.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gould-Rothberg BE, Dipippo VA, Ramseh TM, Gerwein RW;
PI WPI; 2001-355948/37.
XX
XX Screening hepatotoxic agent comprises contacting test cell population
PT expressing RISKMARKER or INJURYMARKER with agent, comparing expression
PT with reference population and identifying difference in expression
PT levels.
XX
XX Disclosure; Page 22-24; 76pp; English.
PS
XX The present invention describes a method of screening a test agent for
CC hepatotoxicity. The method comprises: (a) providing a test cell
CC population comprising a cell capable of expressing one or more nucleic
CC acid sequences selected from the group consisting of RISKMARKER 1-8 and
CC INJURYMARKER 1-10; (b) contacting the test cell population with a test

CC agent; (c) measuring expression of one or more of the nucleic acid
CC sequences in the test cell population; (d) comparing the expression of
CC the nucleic acid sequence in the test cell population to the expression
CC of the nucleic acid sequence in a reference cell population comprising
CC at least one cell whose exposure status to a hepatotoxic agent is known;
CC and (e) identifying a difference in expression levels of the RISKMARKER
CC or INJURYMARKER sequences, if present, in the test cell population and
CC reference cell population. The method is useful for identifying a
CC hepatotoxic agent. The present sequence is given in the exemplification
CC of the present invention

XX
SQ Sequence 5001 BP; 1225 A; 1204 C; 1261 G; 1311 T; 0 U; 0 Other;
XX
Query Match 18.1%; Score 76.4; DB 4; Length 5001;
Best Local Similarity 98.7%; Pred. No. 4.8e-13;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCAGCAGCATTTGAACACTGCACACGCCCATCTGCCAGAGAGCTGTGACAC 400
DB 1 CACAAACCCAGCAGCATTTGAACACTGCACACGCCCATCTGCCAGAGAGCTGTGACAC 60
QY 401 CACTTCGCTACTAGCTA 418
DB 61 CACTTCGCTACTAGCTA 78

RESULT 14
ABN95896
ID ABN95896 standard; DNA; 6128 BP.
XX
AC ABN95896;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #2394 used to diagnose liver cancer.
XX
KM Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumor; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR MPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
XX carcinoma or metastatic liver tumor in a patient, involves detecting the
XX level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 2394; 298bp; English.

CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,

CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WFO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
XX
Query Match 14.4%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.5e-08;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACCACTTATTTGAACACGGGATCTAGACAGCTGCCCTGACATCAT 265
DB 473 CACTAGCAAAACCACTTATTTGAACACTAGCTCTAGCCTGGCGCTGCCATCAT 532
QY 266 TAACCC-----GTGCTGCCAGCAGCCCTTCATTAAGCCCTGGATAGCCACG 315
DB 523 TAACCTCTGTGCAAGTGGCGCGGCTGTGCTTTTAAAGTGGCGGCTGTCTCAGC 592
QY 316 CAGCATGTCTCACTTCCGCGCGCGGAGACACAAACCCGAGCATTGAACACTGC ACACGG 374
DB 593 GAGCATCGGCGCACCGCCATTC-----CATCCGAGGAGCATTCGCCGCGCGCGCG 644
QY 375 CCATCTGCCAGAGAGCTGTGACACCACTTC 406
DB 645 CCACCTCCAGAGGACACTGGCCACCGCTCC 676

RESULT 15
ABV75371
ID ABV75371 standard; DNA; 6128 BP.
XX
AC ABV75371;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human IGFBP-1 gene sequence.
XX
KM Insulin-like growth factor binding protein; IGFBP; cytoablastic; liver;
XX cancer; human; IGFBP-1; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200290580-A1.
XX
FH Key Location/Qualifiers
FT CDS 762..5200
FT FT /tag= a
FT FT /product= "IGFBP-1"
FT FT /note= "insulin-like growth factor binding protein;
FT FT contains introns"
FT FT 762..1110
FT FT /tag= b
FT FT /number= 1
FT FT 1111..2656
FT FT /tag= c
FT FT /number= 1
FT FT 2657..2826
FT FT /tag= d
FT FT /number= 2
FT FT 2827..4040
FT FT /tag= e
FT FT /number= 2
FT FT 4041..4169
FT FT /tag= f
FT FT /number= 3
FT FT 4170..5068
FT FT /tag= g
FT FT /number= 3
FT FT 5069..5197
FT FT /tag= h
FT FT /number= 4
XX
XX

PD 14-NOV-2002.
XX
XX 03-MAY-2002; 2002MO-AU000558.
XX
XX 03-MAY-2001; 2001US-0288441P.
XX
XX (NACA-) NAT CANCER CENT SINGAPORE PTE LTD.
PA (SIGE-) SINGAPORE GEN HOSPITAL PTE LTD.
PA (ARGA/) ARGAMET V P.
XX
PI Huynh TH, Chow PKH, Soo KC;
XX
XX WPI; 2003-103522/09.
DR P-PSDB; ABB82757.
XX
PT Detecting the presence or diagnosing the risk of a liver cancer in a
PT patient comprises detecting aberrant expression of a gene encoding an
PT insulin-like growth factor binding protein.
XX
XX Example; Page 104-108; 142pp; English.
PS
CC The invention relates to detecting the presence or diagnosing the risk of
CC a liver cancer in a patient. The method involves detecting in a
CC biological sample obtained from the patient aberrant expression of a gene
CC encoding an insulin-like growth factor binding protein (IGFBP). The
CC method is useful for detecting the presence or diagnosing the risk of a
CC liver cancer or for screening agents in a patient. The agent is useful
CC for the manufacture of a medicament for treating and/or preventing liver
CC cancer. The present sequence represents a human IGFBP-1 polypeptide
CC encoding genomic DNA (GenBank Accession No. M74587)
XX
SQ Sequence 6128 BP; 1504 A; 1497 C; 1504 G; 1623 T; 0 U; 0 Other;
Query Match 14.4%; Score 60.8; DB 10; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.5e-08;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;
QY 206 CACAAGCAAAACAACTTATTTTGAACACGGGATCCTAGACAGCTGCCCTGACAATCAT 265
DB 473 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGCGGCGCTGCCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATAGGCCCTGGGTATGCGCAGC 315
DB 533 TAACCTCCTGTGCAAGTGCGCGGCTGTGCTTTATTAAGTGTGCGGCTGTGCCAGC 592
QY 316 CAGCATGTGTCCACTGCGCCCGCAGACACAACCCAGCGCATTTGAACACTGC-ACACGG 374
DB 593 GAGCATCGGCGCACCGCCATCC-----CATCCAGCGAGCATGTGCCGCGCGCGCGCG 644
QY 375 CCATCGCCCGAGAGAGCTGTGACCACTTC 406
DB 645 CCACCTCCCGAGAGAGCACTGCGCACCGCTCC 676

Search completed: September 1, 2005, 00:34:50
Job time : 451.543 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 317.549 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51
Sequence: 1 catggggcgcacggggcgcac.....ggactctggcccccagtgtga 51

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	194	6 A57715	A57715 Sequence 1
2	51	100.0	194	6 AR175909	AR175909 Sequence 1
3	51	100.0	13011	6 E14395	E14395 gDNA encodi
4	51	100.0	13011	6 AX827302	AX827302 Sequence
5	51	100.0	13011	10 RNLPKG	X05684 Rat L-PK ge
6	43.2	84.7	2717	10 RATPKRL1	M17088 Rat pyruvat
7	43.2	84.7	231241	2 AC097039	AC097039 Rattus no
8	36.6	71.8	415	9 HSPYKINPR	Z18922 H.sapiens D
9	36.6	71.8	8409	9 HSU47654	U47654 Homo sapien
10	36.6	71.8	13322	9 AY316591	AY316591 Homo sapi
11	36.6	71.8	13325	9 AL713999	AL713999 Human DNA
12	33.6	65.9	192549	2 AC148047	AC148047 Ocolemur
13	32.2	63.1	95718	2 AC147942	AC147942 Ateleserix
14	30.2	59.2	136883	2 AC147944	AC147944 Ateleserix
15	30.2	59.2	153222	2 AC148054	AC148054 Ateleserix
16	30	58.8	13637	2 AC148066	AC148066 Canis fam
17	30	58.8	174734	2 AC148068	AC148068 Canis fam
18	28.8	56.5	178895	2 AC148063	AC148063 Bos tauru
19	25.8	50.6	140869	2 AC148210	AC148210 Macropus

c 20	25.8	50.6	166576	2 AC148209	AC148209 Macropus
c 21	25.2	49.4	158643	2 AC143451	AC143451 Macaca mu
c 22	25.2	49.4	174642	2 AC143448	AC143448 Macaca mu
c 23	25.2	49.4	239210	2 AC143094	AC143094 Macaca mu
c 24	25.2	49.4	349323	1 BX572096	BX572096 Prochiloro
c 25	25	49.0	149618	9 AP000556	AP000556 Homo sapi
c 26	25	49.0	150036	9 AP000557	AP000557 Homo sapi
c 27	25	49.0	166861	2 AC148587	AC148587 Pan trogl
c 28	25	49.0	167076	2 AC136511	AC136511 Pan trogl
c 29	25	49.0	176051	9 AC023490	AC023490 Homo sapi
c 30	25	49.0	212651	9 AC007957	AC007957 Homo sapi
c 31	25	49.0	212651	9 AC007957	AC007957 Homo sapi
c 32	24.6	48.2	173153	9 AC008119	AC008119 Homo sapi
c 33	24.6	48.2	179773	2 AC116041	AC116041 Papio ham
c 34	24.6	48.2	188462	2 AL592301	AL592301 Human DNA
c 35	24.6	48.2	248002	2 AC098668	AC098668 Rattus no
c 36	24.4	47.8	1255	9 HSU79264	HSU79264 Human clone
c 37	24.4	47.8	3086	6 CQ715195	CQ715195 Sequence
c 38	24.4	47.8	3138	6 AR103244	AR103244 Sequence
c 39	24.4	47.8	3138	6 AX409069	AX409069 Sequence
c 40	24.4	47.8	3138	6 BD000107	BD000107 Different
c 41	24.4	47.8	3138	9 HDMZICP	D76435 Homo sapien
c 42	24.4	47.8	102386	4 AC091404	AC091404 Sus scrofa
c 43	24.4	47.8	151453	2 AC118988	AC118988 Sus scrofa
c 44	24.4	47.8	174559	2 AC027060	AC027060 Homo sapi
c 45	24.4	47.8	179714	9 AC092959	AC092959 Homo sapi

ALIGNMENTS

RESULT 1	A57715	Sequence 1 from Patent WO9632489.	194 bp	DNA	linear	PAT 03-MAR-1998
LOCUS	A57715					
DEFINITION	A57715	GI:3713539				
ACCESSION	A57715					
VERSION	A57715.1					
KEYWORDS	unidentified					
SOURCE	unidentified					
ORGANISM	unclassified.					
REFERENCE	1	Chen, R., Doiron, B. and Kahn, A.				
AUTHORS		GLUCOSE-INDUCIBLE RECOMBINANT VIRAL VECTOR				
TITLE		Patent: WO 9632489-A 1 17-OCT-1996;				
JOURNAL		INST NAT SANTE RECH MED (FR)				
COMMENT		Other publication AU 5652396 961030				
		Other publication FR 2732978 961018.				
FEATURES		Location/Qualifiers				
source		1..194				
		/organism="unidentified"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:32644"				
ORIGIN						
Query Match		100.0%; Score 51; DB 6; Length 194;				
Best Local Similarity		100.0%; Pred. No. 2.7e-06;				
Matches	51; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	CATGGGGCAGGGGCACTCCGCGTCTCGACTCTGGCCCCAGTGTGA 51				
Db	11	CATGGGGCAGGGGCACTCCGCGTCTCGACTCTGGCCCCAGTGTGA 61				
LOCUS	AR175909	AR175909	194 bp	DNA	linear	PAT 17-DEC-2001
DEFINITION	Sequence 1 from patent US 6309878.					
ACCESSION	AR175909					
VERSION	AR175909.1	GI:17917208				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					

Unclassified.
REFERENCE 1 (bases 1 to 194)
AUTHORS Chen,R., Doiron,B. and Kahn,A.
TITLE Glucose-inducible recombinant viral vector
JOURNAL Patent: US 6309878-A 1 30-OCT-2001;
FEATURES Location/Qualifiers
source 1..194
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 194;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 51
|||||
11 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 61

Db
11 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 61

RESULT 3
E14395 13011 bp DNA linear PAT 28-JUL-1999
LOCUS gDNA encoding pyruvate kinase.
DEFINITION E14395
ACCESSION E14395.1 GI:5709078
VERSION JP 1997313059-A/14.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 13011)
Murse,M., Murase,J., Iwabuchi,M., Hayakawa,T. and Imamura,J.
INCREASE IN STORED LIPID CONTENT OF PLANT SEED
Patent: JP 1997313059-A 14 09-DEC-1997;
MITSUBISHI CORP, MITSUBISHI CHEM CORP
OS Rattus norvegicus (rat)
PN JP 1997313059-A/14
PD 09-DEC-1997
PE 31-JAN-1997 JP 1997018966
PR 01-FEB-1996 JP 96P 16560
PI MURASE MAKOTO, MURASE JUNKO, IWABUCHI MARI, HAYAKAWA TAKAHIKO,
PI IMAMURA JUN
PC A01H5/00,C07H21/04,C12N5/10,C12N9/12,C12N15/09,C12N5/10, PC
C12N1:91),
PC (C12N9/12,C12R1:91), (C12N15/09,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH source 1..13011
FT 3194..3218
/organism='Rattus norvegicus' FT exon
/number=1
FT intron 3219..3765
/number=1
FT exon 3766..3948
/number=2
FT intron 3949..5916
/number=2
FT exon 5917..6008
/number=3
FT intron 6009..6151
/number=3
FT exon 6152..6283
/number=4
FT intron 6284..6417
/number=4
FT exon 6418..6604
/number=5

FT intron 6605..6920
/number=5
FT exon 6921..7191
/number=6
FT intron 7192..7301
/number=6
FT exon 7302..7452
/number=7
FT intron 7453..7664
/number=7
FT exon 7665..7817
/number=8
FT intron 7818..7910
/number=8
FT exon 7911..8077
/number=9
FT intron 8078..9297
/number=9
FT exon 9298..9479
/number=10
FT intron 9480..10162
/number=10
FT exon 10163..10269
/number=11
FT 5'UTR 1..3193
FT 3'UTR 10270..13011
FT CDS
FT join(3212..3218,3766..3948,5917..6008, FT
6152..6283,
6418..6604,6921..7191,7302..7452, FT
7665..7817,7911..8077,
9298..9479,10163..11594)
/product='Pyruvate kinase'.
FEATURES
source
1..13011
/organism='Rattus norvegicus'
/mol_type='genomic DNA'
/db_xref='taxon:10116'

ORIGIN
Query Match 100.0%; Score 51; DB 6; Length 13011;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 51
|||||
3021 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 3071

Db 3021 CATGGCGCACGGGGCACTCCCGTGTCTCTGACTCTGGCCCCCAGGTGA 3071

RESULT 4
AX827302 13011 bp DNA linear PAT 12-DEC-2003
LOCUS AX827302
DEFINITION Sequence 36 from Patent EP1344834.
ACCESSION AX827302
VERSION AX827302.1 GI:39837391
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1
AUTHORS Boess,F., Suter-Dick,L. and Wolf,D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 36 17-SEP-2003;
F. HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers
source 1..13011
/organism='Rattus norvegicus'
/mol_type='unassigned DNA'
/db_xref='taxon:10116'

ORIGIN

Query Match 100.0%; Score 51; DB 10; Length 13011;
 Best Local Similarity 100.0%; Pred. No. 1.3e-06;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 51
 |||||
 DB 3021 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 3071

RESULT 6
 LOCUS R1P7KRL1 2717 bp DNA linear ROD 27-APR-1993
 DEFINITION Rat pyruvate kinase gene, exons 1 and 2.
 ACCESSION M17088 J03455
 VERSION M17088.1 GI:206208
 KEYWORDS pyruvate kinase.
 SEGMENT 1 of 4
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 2717)
 AUTHORS Noguchi,T., Yamada,K., Inoue,H., Matsuda,T. and Tanaka,T.
 TITLE The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters
 JOURNAL J. Biol. Chem. 262 (29), 14366-14371 (1987)
 MEDLINE 88007696
 PUBMED 3654663

COMMENT Original source text: Rat DNA (library of J. Bonner), clone lambda-LPK30, and reliculocyte, cDNA to mRNA, clones gPRK(1,3).
 Draft entry and printed copy of sequence for [1] kindly provided by T. Noguchi, 09-SEP-1987.

FEATURES
 source Location/Qualifiers

1..2717
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 prim_transcript 925..>2717
 /note="PK mRNA and introns (alt.)"
 prim_transcript 930..>2717
 /note="PK mRNA and introns (alt.)"
 prim_transcript 934..>2717
 /note="PK mRNA and introns (alt.)"
 exon <980..1079
 /note="R-pyruvate kinase"
 intron 1080..2007
 /note="PK intron A"
 exon <1446..1452
 /note="L-pyruvate kinase"
 intron 1453..2007
 /note="PK intron A (alt.)"
 exon 2008..2190
 /number=2
 intron 2191..>2717
 /note="PK intron B"
 ORIGIN 342 bp upstream of PstI site.

Query Match 84.7%; Score 43.2; DB 10; Length 2717;
 Best Local Similarity 93.8%; Pred. No. 0.00059;
 Matches 45; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 GGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 51
 |||||
 DB 1258 GGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGCGCCCGCAGTGA 1305

RESULT 7
 LOCUS AC097039 231241 bp DNA linear HTG 10-MAY-2003
 DEFINITION Rattus norvegicus clone CH230-61E1, *** SEQUENCING IN PROGRESS ***.

ACCESSION AC097039
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 231241)
 AUTHORS Murny,D.,Marie, Metzker,M.,Lee, Abramson,S., Adams,C., Alder,J.,

Allen,C., Allen,H., Alebrooks,S., Amlin,A., Angiano,D.,
 Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Bubay,C., Burch,P., Burrell,K., Calderon,E.,
 Cardenas,V., Carter,K., Carazos,I., Ceaar,H., Centre,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,X., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,D.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denson,S., Derramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Duhdin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,D., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
 Hollins,B., Howell,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Kapachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lounsbury,L., Loussaged,H., Loyado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,K., Martinez,E.,
 Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwokeneme,O., Okwom,G., Olarinuwaogun,A., Pal,S., Parks,K.,
 Patelrnak,S., Paul,H., Perez,A., Perez,L., Plannoch,C.,
 Plapper,F., Polndexter,A., Popovic,D., Primus,E., Pu,L.,L.,
 Puzo,M., Quiroz,J., Rachin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Saverly,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,J., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Syatek,A., Taber,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,Y., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczky,R., Woodson,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
 Weinstock,G., and Gibbs,R.A.

Direct Submission
 TITLE Unpublished
 REFERENCE 2 (bases 1 to 231241)
 AUTHORS Worley,K.C.
 JOURNAL Direct Submission

Submitted (06-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231241)
 Rat Genome Sequencing Consortium.

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:24819342.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHYP

Center clone name: CH230-61E1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 210827 bases at least Q40

Consensus quality: 214055 bases at least Q30

Estimated insert size: 223098; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

- * NOTE: Estimated insert size may differ from sequence length
- * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 1 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submitter.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

1 231241: contig of 231241 bp in length.

FEATURES

source

1. 231241

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-61E1"

misc_feature

1

/note="clone boundary"

clone end:sp6

site:ECORI

misc_feature

end_sequence:BH282405"

755..1528

/note="clone boundary"

clone end:T7

site:ECORI

misc_feature

end_sequence:BH282404"

66116..68518

/note="wgs_end_extension"

clone end:T7"

68569..69982

/note="wgs_end_extension"

clone_end:T7"

misc_feature

clone_end:T7"

ORIGIN

Query Match

Best Local Similarity 84.7%; Score 43.2; DB 2; Length 231241;

Matches 45; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

4 GGGCGACGGGCACTCCGTTCTCTGACTCTGGCCCCCAAGTGA 51

DB

175858 GGGCGACGGGCACTCCGTTCTCTGACTCTGGCCCCCAAGTGA 175905

RESULT 8

HSPYKINPR

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (03-DEC-1992) de Medicis E., Universite de Sherbrooke, de

Biochimie, 12e Avenue Nord, Fleurimont, Quebec, Canada, J1H 5N4

Location/Qualifiers

1..415

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/cell_type="Leukocyte"

1..412

/note="liver pyruvate kinase promoter region"

/citation=[1]

ORIGIN

promoter

Query Match

Best Local Similarity 71.8%; Score 36.6; DB 9; Length 415;

Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db

218 CACGGGACATCCGTTCTCTGACTCTGGCCCCCAAGTGA 51

9 CACGGGACATCCGTTCTCTGACTCTGGCCCCCAAGTGA 51

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (30-JAN-1996) Peter Nuernberg, Institute for Medical

Genetics, Charite Medical School of the Humboldt-University Berlin,

Schumannstr. 20/21, Berlin 10098, Germany

Location/Qualifiers

1..8409

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

```
mrna
      join(<1. .139,1145. .1327,2508. .2599,2696. .2827,2961. .3147,
3510. .3780,3877. .4027,4672. .4824,4919. .5085,6325. .6506,
7572. .7678)
/product="pyruvate kinase PK-R isoenzyme"
join(<1. .139,1145. .1327,2508. .2599,2696. .2827,2961. .3147,
3510. .3780,3877. .4027,4672. .4824,4919. .5085,6325. .6506,
7572. .7678)
/codon_start=1
/product="pyruvate kinase PK-R isoenzyme"
/protein_id="AA92535.1"
/db_xref="GI:1230589"
/translation="HSMVPOQPHTESMSIOENISLQSRWSKGRDLAKSLILIGA
PGGPGYLRASVAQLTQELGTAFPOOQLPAMADTFEHLCLDIDSEPARSTSIATIGPARSVLRKEMIK
IATIGPARSVLRKEMIKAMNTARLNFSGHSHYAESIANREAVESFAGSPIS
YRRVPAIDTKGPEIRTGILQSGPSESVLVGSOVLVTVDPARTGNANTVWDY
NIVRVVPGGRITIDGLISLVQKIGPEGLVTVQENGVLSRKGVLPGAOVDPG
LSEQVDRDLRFVEHGVDFVFAFPAKASDVAVRALGPEGHGKIISKIENHGVK
RDEILVSDGIMVAGDLGIEI PAEKVFLAQRMIGRNLAKGPVCATQMLSEMIT
KRPRPRAETSDVANNVLDGADCI MLGETAKNFPVEAVQMOMAIAREAAVYHRQ
FELRRAPLSDPTEVTAIGAVEAFCACAAIIVLTGGRSAQLLSRYRRAVIA
VTRSAQAARQVHLCRGFPFLYRPEPRAIMADVDRRVQGIESGKRGFLRVGLVI
VTGWRPGSGYTNIMRVLSIS"
join(<478. .553,1145. .1327,2508. .2599,2696. .2827,
2961. .3147,3510. .3780,3877. .4027,4672. .4824,4919. .5085,
6325. .6506,7572. .7678)
/product="pyruvate kinase PK-L isoenzyme"
join(478. .553,1145. .1327,2508. .2599,2696. .2827,2961. .3147,
3510. .3780,3877. .4027,4672. .4824,4919. .5085,6325. .6506,
7572. .7678)
/codon_start=1
/product="pyruvate kinase PK-L isoenzyme"
/protein_id="AAA92536.1"
/db_xref="GI:1230590"
/translation="WELTTQCGAPDRRPREVSGMEGPGYLRASVAQLTQELG
TAFPOOQLPAMADTFEHLCLDIDSEPARSTSIATIGPARSVLRKEMIK
AMNTARLNFSGHSHYAESIANREAVESFAGSPISYRPAIDTKGPEIRTGILQ
GGPSEVLELVGSOVLVTVDPARTGNANTVWDYPNIVRVVPGGRITIDGLISL
VQKIGPEGLVTVQENGVLSRKGVLPGAOVDPGLSEQVDRDLRFVEHGVDFV
ASFARASDVAVRALGPEGHGKIISKIENHGVKRPDEILVSDGIMVAGDLG
EIPAKVFLAQRMIGRNLAKGPVCATQMLSEMITKRPRPRAETSDVANNVLDG
ADCI MLGETAKNFPVEAVQMOMAIAREAAVYHRQFELRRAPLSDPTEVTA
IGAVEAFCACAAIIVLTGGRSAQLLSRYRRAIAVTRSAQAARQVHLCRGFPFL
YRPEPRAIMADVDRRVQGIESGKRGFLRVGLVI VTGWRPGSGYTNIMRVLSIS
"
```

ORIGIN

Query Match 71.8%; Score 36.6; DB 9; Length 8409;
Best Local Similarity 90.7%; Pred. No. 0.068;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGACATCCCGTCTCGACTGACCTGCCCCCATGTA 51
|||
Db 356 CACGGGACATCCCGTCTCGACTGACCTGCCCCCATGTA 398
|||

RESULT 10
AY316591 13322 bp DNA linear PRI 21-JUN-2003
LOCUS AY316591 Homo sapiens pyruvate kinase, liver and RBC (PKLR) gene, complete cds.
DEFINITION
ACCESSION AY316591
VERSION AY316591.1 GI:32140478
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 13322)
Rieder,M.J., Livingston,R.J., Daniels,M.R., Montoya,M.A.,
Chung,M.-W., Miyamoto,K.E., Nguyen,C.P., Nguyen,D.A., Poel,C.L.,
Robertson,F.D., Schackwitz,W.S., Sherwood,J.K., Wiltrak,L.A. and
Nickerson,D.A.

TITLE Direct Submission
JOURNAL Submitted (06-JUN-2003) Genome Sciences, University of Washington,
1705 NE Pacific, Seattle, WA 98195, USA
COMMENT To cite this work please use: NIEHS-SNPs, Environmental Genome
Project, NIEHS ES15478, Department of Genome Sciences, Seattle, WA
(URL: <http://esg.gs.washington.edu>).
FEATURES
source
1. .13322
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .187
/note="Region not scanned for variation"
misc_feature
9. .157
/rpt_family="ERV"
/rpt_type="dispersed
402
/frequency="0.01"
/replace="c"
<525. .11691
/gene="PKLR"
join(<525. .649,1665. .1847,6189. .6280,6377. .6508,
6643. .6829,7193. .7463,7560. .7710,8355. .8507,8602. .8768,
10008. .10189,11267. .11691)
/gene="PKLR"
join(550. .649,1665. .1847,6189. .6280,6377. .6508,6643,
7193. .7463,7560. .7710,8355. .8507,8602. .8768,10008. .10189,
11267. .11373)
/gene="PKLR"
/codon_start=1
/product="pyruvate kinase, liver and RBC"
/protein_id="AAP6927.1"
/db_xref="GI:32140479"
/translation="WSTIOENISLQSRWSKGRDLAKSLILIGAPGPGYLRASV
AQLQDELGTAFPOOQLPAMADTFEHLCLDIDSEPARSTSIATIGPARSV
LRKEMIKAMNTARLNFSGHSHYAESIANREAVESFAGSPISYRPAIDTKGP
EIRTGILQSGPSESVLVGSOVLVTVDPARTGNANTVWDYPNIVRVVPGGRIT
IDGLISLVQKIGPEGLVTVQENGVLSRKGVLPGAOVDPGLSEQVDRDLRFV
EHGVDFVASFARASDVAVRALGPEGHGKIISKIENHGVKRPDEILVSDGIM
VAGDLGIEIPAEVFLAQRMIGRNLAKGPVCATQMLSEMITKRPRPRAETSDV
ANNVLDGADCI MLGETAKNFPVEAVQMOMAIAREAAVYHRQFELRRAPLSD
PTEVTAIGAVEAFCACAAIIVLTGGRSAQLLSRYRRAIAVTRSAQAARQVHL
CRGFPFLYRPEPRAIMADVDRRVQGIESGKRGFLRVGLVI VTGWRPGSGYTN
IMRVLSIS"

repeat_region
1346. .1567
/rpt_family="MIR"
/rpt_type="dispersed
1572
/gene="PKLR"
/frequency="0.01"
/replace="c"
1745
/gene="PKLR"
/frequency="0.01"
/replace="c"
1905
/gene="PKLR"
/frequency="0.02"
/replace="a"
1951. .2139
/rpt_family="L2"
/rpt_type="dispersed
1956
variation
/gene="PKLR"
/frequency="0.03"
/replace="t"
1960
/gene="PKLR"
/frequency="0.42"
/replace="c"
2081. .4840
/gene="PKLR"

repeat_region
variation
variation
misc_feature


```

repeat_region /note="Region not scanned for variation"
                2140..2451
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 2457..2758
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 2759..2810
                /rpt_family="L2"
                /rpt_type=dispersed
repeat_region 3003..3280
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3302..3318
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3319..3611
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3612..3909
                /rpt_family="Alu"
                /rpt_type=dispersed
repeat_region 3920..5300
                /rpt_family="L2"
                /rpt_type=dispersed
variation      5122
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                5162
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                5162
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                5223
                /gene="PKLR"
                /frequency="0.02"
                /replace="c"
                5301..5619
                /rpt_family="Alu"
                /rpt_type=dispersed
variation      5401
                /gene="PKLR"
                /frequency="0.37"
                /replace="a"
                5620..5705
                /rpt_family="L2"
                /rpt_type=dispersed
repeat_region 5698
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                5890..5995
                /gene="PKLR"
                /note="Region not scanned for variation"
                6075
variation      6075
                /gene="PKLR"
                /frequency="0.03"
                /replace="g"
                6290
                /gene="PKLR"
                /frequency="0.02"
                /replace="t"
                6544
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                6559
                /gene="PKLR"
                /frequency="0.40"
                /replace="c"
                7113
                /gene="PKLR"
                /frequency="0.01"

```

```

variation      7566
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                7736
                /gene="PKLR"
                /frequency="0.01"
                /replace="a"
repeat_region 7788..8094
                /rpt_family="Alu"
                /rpt_type=dispersed
variation      7864
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                7898
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
variation      7943
                /gene="PKLR"
                /frequency="0.01"
                /replace="g"
                7943
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
variation      8306
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                8326
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
variation      8352
                /gene="PKLR"
                /frequency="0.01"
                /replace="c"
                8550
                /gene="PKLR"
                /frequency="0.01"

```

```

Query Match      71.8%; Score 36.6; DB 9; Length 13322;
Best Local Similarity 90.7%; Pred. No. 0.063;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      9 CACGGGACATCCCGTGTCTCTGACTCTGCCCCAGTGA 51
Db      867 CACGGGACATCCCGTGTCTCTGACTCTGCCCCAGTGA 909

```

```

RESULT 11
LOCUS      AL713999/C      133525 bp      DNA      linear      PRI 13-MAY-2003
DEFINITION Human DNA sequence from clone RPl1-263K19 on chromosome 1, complete
sequence.
ACCESSION  AL713999
VERSION    AL713999.28  GI:30722340
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 133525)
AUTHORS  Giltner,R.
TITLE    Direct Submission
JOURNAL  Submitted (13-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On May 14, 2003 this sequence version replaced gi:30523877.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chrl

RP11-263K19 is from the library RPCT-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBAC3.6.

FEATURES

source

```
1..133525
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP11-263K19"
    /clone_lib="RPCT-11.1"
```

ORIGIN

Query Match

Best Local Similarity 71.8%; Score 36.6; DB 9; Length 133525;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGGACCTCCCGTTCCTGACCTCGGCCCCCAGCTGTA 51

Db 129358 CACGGGGACCTCCCGTTCCTGACCTCGGCCCCCAGCTGCA 129316

RESULT 12

AC148047/c

LOCUS

DEFINITION

AC148047

ORDERED

AC148047

VERSION

HTG: HTGS PHASE2; HTGS DRAFT.

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 192549)

Antionellis, A., Ayele, K., Benjamin, B., Blakeley, R. W.,

Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,

Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J.,

Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,

Idol, J. R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P.,

Larson, S., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B.,

Marijules, E. H., Mastiello, C., Maskeri, B., Mcowen, J.,

Mullikin, J. C., Paguitan, C., Portnoy, M. E., Prasad, A., Puri, O.,

Reddi-Dugue, N., Schandier, K., Schneider, M. G., Shah, K., Sison, C.,

Stattip, K. D., Thomas, J. W., Thomas, P. J., Tspouri, V., Vogt, J. L.,

Wetherby, K. D., Young, A. and Green, E. D.

NISC Comparative Sequencing Initiative

Unpublished

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

2 (bases 1 to 192549)

Green, E. D.

Direct Submission

Submitted (29-JAN-2004) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 192549)

Green, E. D.

Direct Submission

Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

On Mar 18, 2004 this sequence version replaced gi:41386816.

Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: http://www.nisc.nih.gov

Contact: nisc.zoo@nhgri.nih.gov

Center project name: fxd

Center clone name: 360P11

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 190022 bases at least Q40

Consensus quality: 191709 bases at least Q30

Consensus quality: 191709 bases at least Q20

Insert size: 149000; agarose-1p

Insert size: 191949; sum-of-contigs

Quality coverage: 11.55x in Q20 bases; agarose-1p

Quality coverage: 8.97x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 7 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

This sequence will be replaced

by the finished sequence as soon as it is available and

the accession number will be preserved.

1 38219: contig of 38219 bp in length

38220 38319: gap of unknown length

38320 77219: contig of 38900 bp in length

77220 77319: gap of unknown length

77320 114115: contig of 36796 bp in length

114116 114215: gap of unknown length

114216 126274: contig of 12059 bp in length

126275 126374: gap of unknown length

126375 152697: contig of 26323 bp in length

152698 152977: gap of unknown length

152978 165906: contig of 14109 bp in length

165907 167007: gap of unknown length

167007 192549: contig of 25543 bp in length.

FEATURES

source

```
1..192549
    /organism="Otolemur garnettii"
    /mol_type="genomic DNA"
    /db_xref="taxon:30611"
    /clone="CH256-360P11"
    /clone_lib="CH256"
```

```

misc_feature      1.38219 /note="BAC resource: http://bacpac.chori.org/"
                  /note="assembly_fragment"
                  clone_end:SP6
                  vector_side:left"
misc_feature      1..30749 /note="clone overlaps with GenBank Accession Number
AC148125 clone CH256-404N19 (center project name fxe)"
misc_feature      38320..77219 /note="assembly_fragment"
misc_feature      77320..114115 /note="assembly_fragment"
misc_feature      114216..126274 /note="assembly_fragment"
misc_feature      126375..152697 /note="assembly_fragment"
misc_feature      152798..166906 /note="assembly_fragment"
misc_feature      167007..192549 /note="assembly_fragment"
misc_feature      /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"

ORIGIN
Query Match      65.9%; Score 33.6; DB 2; Length 192549;
Beef Local Similarity 81.2%; Pred. No. 0.36;
Matches 39; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY      4 GGGCGCAGCGGCGACTCCCGTGGTTCCTGAGACTGCGCCCGAGTGA 51
          |||||
Db      94043 GGAGCGACGGCGCATTCCTCGTGTCTCGAGACTTGGCCCGTGGCTTA 93996

RESULT 13
AC147942/c
LOCUS      AC147942/c
DEFINITION AC147942      95718 bp. DNA. linear. HTG 18-MAR-2004
Carollia peripicillata clone 47203, WORKING DRAFT SEQUENCE, 5
ordered pieces.
ACCESSION      AC147942
VERSION      AC147942.2 GI:45544627
KEYWORDS      HTG, HTGS PHASE2; HTGS_DRAFT.
SOURCE      Carollia peripicillata (Seba's short-tailed bat)
ORGANISM      Carollia peripicillata
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
              Carollinae; Carollia.
REFERENCE      1 (bases 1 to 95718)
AUTHORS      Antonellis,A., Ayele,K., Benjamin,B., Blakeley,R.W.,
              Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
              Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
              Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
              Idol,J.R., Jones,C., Karling,E., Kim,H., Kwong,P., Latic,P.,
              Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,
              Marsulles,E.H., Masello,C., Maskeri,B., McDowell,J.,
              Mullikin,J.C., Paguirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
              Reddi-Dugue,N., Schandier,K., Schneider,M.G., Shan,K., Sison,C.,
              Statikipop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
              Wetterly,K.D., Young,A. and Green,E.D.
              NISC Comparative Sequencing Initiative
              Unpublished
              2 (bases 1 to 95718)
              Green,E.D.
              Direct Submission
              Submitted (22-JAN-2004) NIH Intramural Sequencing Center, 8717
              Grovmont Circle, Gaithersburg, MD 20877, USA
              3 (bases 1 to 95718)
              Green,E.D.
              Direct Submission
              Submitted (18-MAR-2004) NIH Intramural Sequencing Center, 8717
              Grovmont Circle, Gaithersburg, MD 20877, USA
              On Mar 18, 2004 this sequence version replaced gi:41058168.
              ----- Genome Center
              Center: NIH Intramural Sequencing Center

```

```

Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: eju
Center clone name: 472003

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 94695 bases at least Q40
Consensus quality: 95070 bases at least Q40
Consensus quality: 95243 bases at least Q20
Insert size: 91000; agarose-fp
Insert size: 95318; sum-of-contigs
Quality coverage: 11.68x in Q20 bases; agarose-fp
Quality coverage: 11.15x in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1      18808: contig of 18808 bp in length
*      18909: gap of unknown length
*      18909: contig of 3211 bp in length
*      22120: gap of unknown length
*      22220: contig of 9531 bp in length
*      31751: gap of unknown length
*      31851: contig of 21213 bp in length
*      53063: gap of unknown length
*      53164: 95718: contig of 42555 bp in length.
*
* Location/Qualifiers
*   1..95718
*     /organism="Carollia peripicillata"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:40233"
*     /clone="47203"
*     /clone_l1b="The Green lab"
*     /note="egreen@nhgri.nih.gov"
*   1..18808
*     /note="assembly_fragment"
*     clone_end:SP6
*     vector_side:left"
*   1..8485
*     /note="clone overlaps with GenBank Accession Number
AC148052 clone 593F8 (center project name gda)"
18909..22119
/note="assembly_fragment"
22220..31750
/note="assembly_fragment"
31851..53063
/note="assembly_fragment"
53164..95718
/note="assembly_fragment"
clone_end:T7

```


misc_feature 139917..148451 /note="assembly_fragment"
misc_feature 148552..153222 /note="assembly_fragment"
clone_end:SP6
vector_side:right"

ORIGIN

Query Match 59.2%; Score 30.2; DB 2; Length 153222;
Best Local Similarity 81.4%; Pred.No. 4.9;
Matches 35; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 2 ATGGGGGACACGGGGGACATCCCGTGTCTCTGGACTCTGGCCCC 44
Db 68059 ATGGGGCCATGGGACAGTCCCATGTCTCTCTGGACTCTGGCCCC 68017

Search completed: September 1, 2005, 03:26:37
Job time : 327.549 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 387.736 Seconds
(without alignments)
5006.706 Million cell updates/sec

Title: US-09-972-916b-1

Perfect score: 51
Sequence: 1 catggggcgcacggggcacc.....ggaccctggcccccctgta 51

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsb1:*
9: gb_gsb2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	84.3	659	8	AZ840793 2M0138D02
2	28.8	56.5	406	4	BM286990 527229 MA
3	27.6	54.1	827	2	BF105536 601822455
4	26.4	51.8	907	2	BE867308 601442381
5	26	51.0	765	1	AL528750 1A528750
6	25.8	50.6	379	9	CE316355 tigr-g88
7	25.8	50.6	672	9	CE450351 tigr-g88
8	25.8	50.2	739	7	CN360924 170004706
9	25.6	50.2	225	6	CD630056 56027932H
10	25.6	50.2	229	6	CD630055 56027932H
11	25.6	50.2	1021	5	BO056161 AGENCOURT
12	25.2	49.4	472	2	AM922227 DGI_17_E0
13	25	49.0	259	8	CC453282 ZMMBC034
14	25	49.0	334	4	BI006710 RCS-RT005
15	25	49.0	403	9	CE744237 tigr-g88
16	25	49.0	481	2	AM577607 PM2-BT055
17	25	49.0	966	5	BU186634 AGENCOURT
18	25	49.0	969	4	BG482095 602527983
19	24.6	48.2	865	8	BZ404685 CGAAM77TC
20	24.6	48.2	1193	5	BQ438639 AGENCOURT
21	24.4	47.8	416	9	CE608928 tigr-g88
22	24.4	47.8	416	9	CE565764 tigr-g88
23	24.4	47.8	460	2	BF189053 234314 MA
24	24.4	47.8	844	1	AU079412 AU079412

C 25	24.4	47.8	884	2	BF101581 601753396
26	24.4	47.8	899	9	CC972844 ZUACY29TH
27	24.4	47.8	913	5	BQ940337 AGENCOURT
C 28	24.2	47.5	265	9	CE418242 tigr-g88
29	24.2	47.5	267	9	CG481325 OST13706
C 30	24.2	47.5	377	6	CB177707 tigr-g88
C 31	24.2	47.5	413	7	CO739485 S1LB06a23
32	24.2	47.5	442	9	CE200317 tigr-g88
C 33	24.2	47.5	469	9	CE699841 tigr-g88
34	24.2	47.5	518	7	CN669363 A0878H10
35	24.2	47.5	521	5	BM944811 UI-M-EH0D
36	24.2	47.5	523	9	CE338459 tigr-g88
37	24.2	47.5	528	7	CN670899 A0898H12
38	24.2	47.5	539	5	BM943713 UI-M-EH0D
39	24.2	47.5	541	7	CE627560 tigr-g88
40	24.2	47.5	552	9	CE668350 tigr-g88
41	24.2	47.5	569	5	BM944273 UI-M-EH0D
42	24.2	47.5	572	9	CE075799 tigr-g88
43	24.2	47.5	594	7	CE619547 tigr-g88
44	24.2	47.5	595	9	CE123626 tigr-g88
45	24.2	47.5	610	9	CE256374 tigr-g88

ALIGNMENTS

RESULT 1
AZ840793 659 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0138D02R Mouse 10kb plasmid U08C1M library mus musculus genomic
LOCUS AZ840793
clone U08C2M0138D02 R, genomic survey sequence.

ACCESSION AZ840793
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niedermaier, A., and Wright, D., Weiss, R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
COMMENT
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

EMAIL: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0138 row: D column: 02
Seq primer: CACACAGCAACAGCTATGACC
Class: plasmid ends
High quality sequence atp: 659.
Location/Qualifiers

FEATURES

source
1..659
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U08C2M0138D02"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid U08C1M library"
/note="Vector: pMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 84.3%; Score 43; DB 8; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.00051;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTTCCTGACTGCGCCCACTGTA 51
|||||

Db 434 CACGGGCACTCCCGTTCCTGACTGCGCCCACTGTA 476

RESULT 2
LOCUS BM286990 406 bp mRNA linear EST 28-DEC-2001
DEFINITION 527829 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM286990
VERSION BM286990.1 GI:17996016
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 406)

REFERENCE
AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Petrea,G., Holt,I., Karaycheva,S., Liang,F.,
Quackenbush,J. and Keefe,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)

JOURNAL MEDLINE PubMed
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCACGACG
Plate: 135 row: G column: 20
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers

FEATURES
source 1..406
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3BOV"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal
longissimus muscle."

ORIGIN

Query Match 56.5%; Score 28.8; DB 4; Length 406;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 36; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 GGGGCGACGGGCACTCCCGTTCCTGACTGCGCCCACTGTA 51
|||||

Db 124 GGAGCGAAGGCGCACTCCCATGACTGACTGCGCCCTGTGCA 171

RESULT 3
LOCUS BF105536 827 bp mRNA linear EST 19-OCT-2000
DEFINITION 601822455R1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4045620 3',
mRNA sequence.
ACCESSION BF105536
VERSION BF105536.1 GI:10888062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 827)

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

JOURNAL Contact: Robert Strusberg, Ph.D.

COMMENT Email: csapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA library Preparation: CLONTECH Laboratories, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L1CM871 row: m column: 13
High quality sequence start: 28
High quality sequence stop: 34.
Location/Qualifiers

FEATURES
source 1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="IMAGE:4045620"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 75"
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgcc); Site 2: SfiI (ggccatratgccc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGCGCCATTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.65
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 54.1%; Score 27.6; DB 2; Length 827;
Best Local Similarity 78.6%; Pred. No. 78;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 GGGGCGACGGGCACTCCCGTTCCTGACTGCGCCCACTGTA 45
|||||

Db 404 GGGGCGAAGGCGCACTCCCGGCTTCGTGACGCTCGCCCC 363

RESULT 4
LOCUS BE867308 907 bp mRNA linear EST 20-OCT-2000
BE867308


```

DEFINITION      601442381F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3846429 5',
                  mRNA sequence.
ACCESSION       BB667308
VERSION         BB667308.1 GI:10316084
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE       NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS         National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE           Unpublished (1999)
JOURNAL         Unpublished (1999)
COMMENT         Contact: Robert Strausberg, Ph.D.
                 Email: cgabbs@mail.nih.gov
                 Tissue Procurement: ATCC
                 cDNA Library Preparation: Life Technologies, Inc.
                 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
                 DNA Sequencing by: Incyte Genomics, Inc.
                 Clone distribution: MGC clone distribution information can be
                 found through the I.M.A.G.E. Consortium/LIML at:
                 http://image.llnl.gov
                 Plate: LLM9559 row: a column: 22
                 High quality sequence stop: 563.
FEATURES        Location/Qualifiers
                 1..907
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:3846429"
                 /tissue_type="adenocarcinoma"
                 /lab_host="DH10B (phage-resistant)"
                 /clone_1lb="NIH_MGC_65"
                 /note="Organ: Colon; Vector: PCMV-SPORT6; Site_1: NotI;
                 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                 Average insert size 1.8 kb. Library constructed by Life
                 Technologies."
ORIGIN
Query Match      51.8%; Score 26.4; DB 2; Length 907;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 33; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY
4 GGGCGACGGGACATCCCGTGTCTCTGACTCTGCCCCCAG 47
|||
777 GGCTGCACGGGCGCCCTGGCGGAGACTGTCTGTGCGACGAG 820
|||
RESULT 5
AL528750      765 bp      mRNA      linear      EST 24-MAR-2004
LOCUS         AL528750 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION    cDNA clone CSODD001YN07 3-PRIME, mRNA sequence.
ACCESSION     AL528750
VERSION       AL528750.3 GI:45703809
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 765)
               L1.W.B., Gruber,C., Jessee,J. and Polayer,D.
               Full-length cDNA libraries and normalization
               Unpublished (2001)
               On Feb 13, 2001 this sequence version replaced gi:31066600.
               Contact: Genoscope
               Genoscope - Centre National de Sequenage
               2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
               Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
               1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
               end enriched, double-strand cDNA was digested with Not I and cloned
               into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library
               was normalized. Library was constructed by Life Technologies, a

```

division of Invitrogen. This sequence belongs to sequence cluster 6514.r and it belongs to a clone representative of this cluster. For more information about this cluster and the virtual cDNA, see <http://www.genoscope.cns.fr/cdna?c=CSODD001CG04NP1&c=6514.r>.

```

FEATURES        Location/Qualifiers
                 1..765
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="CSODD001YN07"
                 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
                 /clone_1lb="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
                 /note="1st strand cDNA was primed with a NotI-oligo(dT)
                 primer. Five prime end enriched, double-strand cDNA was
                 digested with Not I and EcoR V sites of the PCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match      51.0%; Score 26; DB 1; Length 765;
Best Local Similarity 76.2%; Pred. No. 2.7e+02;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY
1 CATGGCGACGGGACATCCCGTGTCTCTGACTCTGGCC 42
|||
445 CCTGGGCACACTGTGCATCTCCCGAATCCTGAGACTTGGCC 486
|||
RESULT 6
CE316355      379 bp      DNA      linear      GSS 26-SEP-2003
LOCUS         CE316355/c
DEFINITION    tigr-gss-dog-17000361087224 dog library Canis familiaris genomic.
ACCESSION     CE316355
VERSION       CE316355.1 GI:36119120
KEYWORDS      GSS.
SOURCE        Canis familiaris (dog)
ORGANISM      Canis familiaris
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
               1 (bases 1 to 379)
               Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
               Rueck,D.B., Delcher,A.L., Pop,M., Wang,W., Frazer,C.M. and
               Venter,J.C.
               The dog genome: survey sequencing and comparative analysis
               Science 301 (5641), 1898-1903 (2003)
               22875432
               MEDLINE
               PUBMED
               14512627
               Contact: Kirkness EF
               The Institute for Genomic Research
               Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
               Rockville, MD 20850, USA
               Tel: 301-838-0200
               Fax: 301-838-0208
               Email: ekirknes@tigr.org
               Classes: shotgun.
FEATURES        Location/Qualifiers
                 1..379
                 /organism="Canis familiaris"
                 /mol_type="genomic DNA"
                 /strain="Standard Poodle"
                 /db_xref="taxon:9615"
                 /clone_1lb="Dog Library"
                 /note="Site 1: BstXI; Libraries were prepared from
                 peripheral blood"
ORIGIN
Query Match      50.6%; Score 25.8; DB 9; Length 379;
Best Local Similarity 81.1%; Pred. No. 3e+02;
Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY
1 CATGGCGACGGGACATCCCGTGTCTCTGAGACTC 37
|||

```

Db 346 CATGGGACAGTGGGCACTCCGCGCTCCCTGGGCTC 310

RESULT 7
LOCUS CE450351 672 bp DNA linear GSS 27-SEP-2003
DEFINITION tigr-gss-dog-17000319093723 Dog Library Canis familiaris genomic,
genomic survey sequence.

ACCESSION CE450351

VERSION CE450351.1 GI:36740297

KEYWORDS GSS.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Bakeryota; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 672)
Kirkness,E.F., Batina,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Frazer,C.M. and
Venter,J.C.

COMMENT The dog genome: survey sequencing and comparative analysis

TITLE Science 301 (5641), 1898-1903 (2003)

JOURNAL 22875432

MEDLINE 14512627

PUBMED Contact: Kirkness EF

FEATURES The Institute for Genomic Research

source Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Classes: shotgun.

Location/Qualifiers

1..672

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strain="Standard Poodle"

/db_xref="taxon:9615"

/clone_1ib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 50.6%; Score 25.8; DB 9; Length 672;

Best Local Similarity 81.1%; Pred. No. 3.1e+02;

Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 CATGGGCGGACGCGGCACTCCGCGCTCCCTGGGCTC 37

Db 553 CATGGGCGGACGCGGCACTCCGCGCTCCCTGGGCTC 589

RESULT 8

LOCUS CN360924 739 bp mRNA linear EST 16-MAY-2004

DEFINITION 17000470668057 GRN_EB Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN360924

VERSION CN360924.1 GI:47360858

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 739)
Brandenberger,R., Wei,H., Zhang,S., lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Pang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J. and Stanton,L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 739 Std Error: 0.00.
Location/Qualifiers

FEATURES

source

1..739

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/issue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"

/clone_1ib="GRN_EB"

/note="Oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Query Match 50.6%; Score 25.8; DB 7; Length 739;

Best Local Similarity 81.1%; Pred. No. 3.1e+02;

Matches 30; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 8 GCACGGGCACTCCGCGTCTCTGAGCTTGCCGCC 44

Db 231 GCACGGGCACTCCGCGTCTCTGAGCTTGCCGCC 195

RESULT 9

LOCUS CD630056 225 bp mRNA linear EST 12-JAN-2004

DEFINITION 56027932J1 FLP Homo sapiens cDNA, mRNA sequence.

ACCESSION CD630056

VERSION CD630056.1 GI:40278322

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 225)
Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES

source

1..225

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_1ib="FLP"

/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 50.2%; Score 25.6; DB 6; Length 225;

Best Local Similarity 77.5%; Pred. No. 3.4e+02;

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 8 GCACGGGCACTCCGCGTCTCTGAGCTTGCCGCC 47

Db 85 GCACGTGACCAACCCGGGAGCTTGAGCTTGATCAG 124

RESULT 10

LOCUS CD630055 229 bp mRNA linear EST 12-JAN-2004

DEFINITION 56027932H1 FLP Homo sapiens cDNA, mRNA sequence.

ACCESSION CD630055

VERSION CD630055.1 GI:40278321

KEYWORDS EST.

[illegible]

VERSION CC453282.1 GI:31000845
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 259)
REFERENCE Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Rouzard,K., Fuks,G., Yu,Y., Wang,R. and Messing,J.
AUTHORS Sequencing of the maize genome at PGIR (2003b)
TITLE Unpublished (2003)
JOURNAL Contact: Bharti,A.K.
COMMENT Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 83.
Location/Qualifiers
1..259
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMB0344B20"
/lab_host="E. coli DH10B"
/clone_lib="ZMMB03"
/note="Vector: pTARBAC1.3; Site_1: BamHI; site_2: BamHI"

ORIGIN
Query Match 49.0%; Score 25; DB 8; Length 259;
Best Local Similarity 69.4%; Pred. No. 5.4e+02;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 3 TGGGGGCGACGGGCACTCCCGTCTCTGGACCTGCGCCCACTGTA 51
123 TCGGCTACGGGAGAGCCCGATGATGCTGAGCTGCTCCCGCGTA 75

Db
RESULT 14
LOCUS BI006710/c 334 bp mRNA linear EST 13-JUN-2001
DEFINITION RCS-RT0053-110101-011-H09 RT0053 Homo sapiens CDNA, mRNA sequence.
ACCESSION BI006710
VERSION BI006710.1 GI:14410781
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 334)
REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
AUTHORS Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922

JOURNAL
MEDLINE
PUBMED
COMMENT
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
SOURCE
FEATURES
location/Qualifiers
1..403
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN
Query Match 49.0%; Score 25; DB 9; Length 403;

Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RCS&t2=RCS-RT0053-110101-011-H09&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 334.
Location/Qualifiers
1..334
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="RT0053"
/note="Organ: kidney tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 49.0%; Score 25; DB 4; Length 334;
Best Local Similarity 75.6%; Pred. No. 5.5e+02;
Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 8 GCACGGGGCACTCCCGTGTCTCTGGACCTGCGCCCACTG 48
Db 333 GCACGGGGCACTCCCGTGTCTCTGGACCTGCGCCCACTG 293

RESULT 15
LOCUS CE744237/c 403 bp DNA linear GSS 30-SEP-2003
DEFINITION tigr-gss-dog-17000369534847 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE744237
VERSION CE744237.1 GI:37084584
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 403)
REFERENCE Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
AUTHORS The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES
location/Qualifiers
1..403
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN
Query Match 49.0%; Score 25; DB 9; Length 403;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 12:52:03 ; Search time 54.0797 Seconds
(without alignments)
5582.626 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51

Sequence: 1 catggcgccagcgagccacac.....gagactctgccccacgtgta 51

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002as:*
6: geneseqn2002bs:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2003ds:*
11: geneseqn2004as:*
12: geneseqn2004bs:*
13: geneseqn2004cs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	100.0	51	ABX15371	Abx15371 Rat glucoc
2	51	100.0	194	ABT43002	Abt43002 Rat type
3	51	100.0	270	ABX15373	Abx15373 Rat insul
4	51	100.0	321	ABX15374	Abx15374 Rat insul
5	51	100.0	372	ABX15375	Abx15375 Rat insul
6	51	100.0	423	ABX15376	Abx15376 Rat insul
7	51	100.0	13011	2 AAT96631	Aat96631 CDNA enco
8	51	100.0	13011	10 ABT42448	Abt42448 Toxicity
9	51	100.0	13011	12 ADP72914	Adp72914 Renal tox
10	48	94.1	48	ABX15379	Abx15379 Rat liver
11	48	94.1	51	ABX15380	Abx15380 Rat liver
12	30	58.8	30	ABE83565	ABE83565 ChREBP pu
13	27	52.9	27	ABE83567	ABE83567 L-pyruvat
14	27	52.9	27	ABE83566	ABE83566 L-pyruvat
15	26	51.0	26	ABE83614	ABE83614 ChREBP ge
16	24.6	48.2	55001	12 ADG88848	Adg88848 Human Not
17	24.6	48.2	55001	12 ADH74825	Adh74825 Human Not
18	24.4	47.8	3138	3 AAA62683	Aaa62683 Human Zic
19	24.4	47.8	3138	6 ABN95218	Abn95218 Gene #171
20	24.4	47.8	3138	12 ADN04235	Adn04235 Antispori

ALIGNMENTS

21	24.4	47.8	3138	12	ADQ17359	Adq17359 Human bof
22	24.4	47.8	3138	13	ADR25048	Adr25048 Breast ca
23	24.4	47.8	3511	12	ADQ22099	Adq22099 Human bof
24	24.4	47.1	24	2	AAT43003	Aat43003 Rat type
25	24	47.1	209	12	ACH93032	Ach93032 Human gen
26	24	47.1	525	12	ACH79332	Ach79332 Human gen
27	24	47.1	712	4	AAK77580	Aak77580 Human imm
28	24	47.1	840	6	AAS98133	Aas98133 Human DNA
29	24	47.1	1188	8	ABV73365	Abv73365 Human TGR
30	24	47.1	1191	6	ABE42892	AbE42892 Human GPC
31	24	47.1	1191	12	ADO29943	Ado29943 Human MAR
32	24	47.1	1227	6	AAI68793	Aai68793 Human MAR
33	24	47.1	1311	9	ACC84331	Acc84331 Human hp1
34	24	47.1	1416	2	AAK77579	Aak77579 Human imm
35	24	47.1	1498	2	AAV38513	Aav38513 Human 7-c
36	24	47.1	1546	6	AAS98070	Aas98070 Human DNA
37	24	47.1	1546	8	ABZ42551	Abz42551 Human G P
38	24	47.1	1546	10	ABT17029	Abt17029 Human MP2
39	24	47.1	1579	4	AAI99548	Aai99548 Human exp
40	24	47.1	1595	4	AAH78712	Aah78712 Human EX3
41	24	47.1	1905	10	ADF70603	Adf70603 Orphan re
42	24	47.1	2046	4	AAS28996	Aas28996 Genomic s
43	24	47.1	2046	4	AAS30237	Aas30237 DNA encod
44	24	47.1	2046	4	AAS35064	Aas35064 DNA #14 e
45	24	47.1	2046	4	AAK77581	Aak77581 Human imm

RESULT 1

ABX15371
ID ABX15371 standard; DNA, 51 BP.

ABX15371;

17-APR-2003 (first entry)

Rat glucose response element (GIRE) DNA.

Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
insulin-sensitive element; ISSE; basal promoter; hypoglycaemia; insulin;
insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoadicosis;
hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
intracellular hormone receptor; insulin regulator construct; anabolic.

Rattus norvegicus.

US2002107198-A1.

08-AUG-2002.

10-OCT-2001; 2001US-00972916.

11-OCT-2000; 2000US-0239113P.

(THUL/) THULE P M.

Thule PM;

WPI; 2002-674190/72.

New insulin regulator cassette, useful e.g. for treating diabetes,
provides specific, glucose-inducible transgenic expression of insulin in
liver cells.

Claim 7; Page 13; 37pp; English.

The invention relates to an insulin regulator construct comprising a
glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
promoter and an insulin-sensitive element (ISSE) of an insulin-like growth

CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat glucose response element of the invention
XX
SQ Sequence 51 BP; 6 A; 18 C; 17 G; 10 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 51
Db 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 51
RESULT 2
AAT43002
ID AAT43002 standard; DNA; 194 BP.
XX
AC AAT43002;
XX
DT 16-JUL-1997 (first entry)
XX
DE Rat type I pyruvate kinase promoter region.
XX
KW Glucose-inducible; rat; pyruvate kinase type I; L-PK; promoter;
KW hyperglycaemia; diabetes; gene therapy; viral vector; defective virus;
KW ss.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Promoter 16..39
FT /tag= a
FT /label= L4_element
FT /note= "able to interact with MLTF/USF factor"
FT promoter 40..58
FT /tag= b
FT /label= L3_element
FT /note= "able to interact with HNF 4 factor"
XX
PN WO9632489-A1.
XX
PD 17-OCT-1996.
XX
PF 12-APR-1996; 96WO-FR000560.
XX
PR 14-APR-1995; 95FR-00004558.
XX
PA (INRM) INST NAT SANTE & RECH MEDICALE.
XX
PI Chen R, Doiron B, Kahn A;
XX
DR WPI; 1996-477139/47.
XX
PT Recombinant defective virus contg. heterologous gene under control of
PT glucose inducible promoter - and infected mammalian cells, esp. for gene
PT therapy of diabetes.
XX
PS Claim 4; Page 21; 41pp; French.

XX
CC A new defective recombinant virus contains at least one heterologous gene
CC under control of an expression signal inducible by glucose or its
CC analogues. Preferably, the expression signal is derived from the 183 bp
CC region located 5' of the type I pyruvate kinase (L-PK) gene, especially
CC having the present sequence. Expression of the recombinant virus can be
CC controlled by glucose levels, but insulin production can be stopped
CC immediately by administering glucagon. This avoids the risk of
CC hypoglycaemia caused by excessive insulin production. Mammalian cells
CC infected by the virus can be implanted into liver, spleen, pancreas or
CC intestine, to provide insulin secretion at predetermined sites. The virus
CC is useful in gene therapy to treat and/or prevent diseases associated
CC with hyperglycaemia, particularly diabetes, but more generally can be
CC used to provide controlled expression of a wide variety of proteins
XX
SQ Sequence 194 BP; 50 A; 59 C; 51 G; 34 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 2; Length 194;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 51
Db 1 CATGGCGCAGCGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 61
RESULT 3
ABX15373
ID ABX15373 standard; DNA; 270 BP.
XX
AC ABX15373;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #1.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ds; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PF New insulin regulator cassette, useful e.g. for treating diabetes,
PF provides specific, glucose-inducible transgenic expression of insulin in
PF liver cells.
XX
PS Claim 9; Page 13-14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to

CC increase fat catabolism and to retard protein catabolism. The construct
CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 270 BP; 63 A; 95 C; 65 G; 47 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 270;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGAGACTCTGAGCCGCCAGTGTGA 51
DB 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGAGACTCTGAGCCGCCAGTGTGA 51
RESULT 4
ABX15374/c
ID ABX15374 standard; DNA; 321 BP.
XX
AC ABX15374;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #2.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct

CC provides stimulation of insulin expression by glucose (but not other
CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
CC combination of these effects can maintain nearly euglycaemic conditions
CC in diabetics during short-term fasting, large carbohydrate loads or when
CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
CC thus inhibiting the long-term complications of diabetes. The properties
CC of the construct are essentially specific for hepatocytes and well-
CC differentiated hepatoma lines and insulin expression in these cells may
CC have effects additional to those provided by secreted insulin, e.g.
CC inhibition of cellular protein degradation, and inhibition, or
CC stimulation of other intracellular hormone receptors. This sequence
CC represents a rat insulin regulator construct of the invention
XX
SQ Sequence 321 BP; 77 A; 111 C; 84 G; 49 T; 0 U; 0 Other;
Query Match 100.0%; Score 51; DB 6; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.5e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGGGCGCACGGGCACTCCCGTGTCTCTGAGACTCTGAGCCGCCAGTGTGA 51
DB 102 CATGGGCGCACGGGCACTCCCGTGTCTCTGAGACTCTGAGCCGCCAGTGTGA 52
RESULT 5
ABX15375/c
ID ABX15375 standard; DNA; 372 BP.
XX
AC ABX15375;
XX
DT 17-APR-2003 (first entry)
XX
DE Rat insulin regulator construct DNA #3.
XX
KW Glucose response element; GIRE; liver pyruvate kinase; LPK; de; IGFBP-1;
KW insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
KW intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
OS Synthetic.
XX
PN US2002107198-A1.
XX
PD 08-AUG-2002.
XX
PF 10-OCT-2001; 2001US-00972916.
XX
PR 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Claim 9; Page 14; 37pp; English.
XX
CC The invention relates to an insulin regulator construct comprising a
CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
CC to treat or prevent diabetic complications, to regulate insulin
CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
CC increase fat catabolism and to retard protein catabolism. The construct

CC carbohydrate) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

Sequence 372 BP; 87 A; 128 C; 102 G; 55 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 372;
 Best Local Similarity 100.0%; Pred. No. 4.5e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGGCGCAGGGGCACTCCCGTGGTTCCTGACTCTGACCCCAAGTGA 51
 |||||
 Db 153 CATGGGCGCAGGGGCACTCCCGTGGTTCCTGACTCTGACCCCAAGTGA 103

RESULT 6
 ABX15376
 ID ABX15376 standard; DNA; 423 BP.
 XX
 AC ABX15376;
 XX
 DT 17-APR-2003 (first entry)
 XX
 DE Rat insulin regulator construct DNA #4.
 XX
 KM Glucose response element; GIRE; liver pyruvate kinase, LPK; ds; IGFPP-1;
 KM insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
 KM insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KM fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KM glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KM hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KM intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 OS Synthetic.
 OS
 PN US2002107198-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL/) THULE P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a

CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a rat insulin regulator construct of the invention
 XX

Sequence 423 BP; 81 A; 149 C; 116 G; 77 T; 0 U; 0 Other;

Query Match 100.0%; Score 51; DB 6; Length 423;
 Best Local Similarity 100.0%; Pred. No. 4.6e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGGGCGCAGGGGCACTCCCGTGGTTCCTGACTCTGACCCCAAGTGA 51
 |||||
 Db 1 CATGGGCGCAGGGGCACTCCCGTGGTTCCTGACTCTGACCCCAAGTGA 51

RESULT 7
 AAT96631
 ID AAT96631 standard; DNA; 13011 BP.
 XX
 AC AAT96631;
 XX
 DT 27-FEB-1998 (first entry)
 XX
 DE cDNA encoding rat cytosolic pyruvate kinase.
 XX
 KM Rattus norvegicus; rat; cytosolic pyruvate kinase; storage lipid; seed;
 KM storage protein; oilseed plant; ds.
 XX
 OS Rattus norvegicus.
 OS
 PH Key location/Qualifiers
 FT exon 1394..3218
 FT /tag= b
 FT /number= 1
 FT CDS 3212..10269
 FT /tag= a
 FT /note= "contains introns"
 FT intron 3219..3765
 FT /tag= c
 FT /number= 1
 FT exon 3766..3948
 FT /tag= d
 FT /number= 2
 FT intron 3949..5916
 FT /tag= e
 FT /number= 2
 FT exon 5917..6008
 FT /tag= f
 FT /number= 3
 FT intron 6009..6151
 FT /tag= g
 FT /number= 3
 FT exon 6152..6283
 FT /tag= h
 FT /number= 4
 FT intron 6284..6417
 FT /tag= i
 FT /number= 4
 FT exon 6418..6604
 FT /tag= j
 FT /number= 5
 FT intron 6605..6920
 FT /tag= k
 FT /number= 5
 FT exon 6921..7191
 FT /tag= l

```
FT      /number= 6
FT      7191. .7301
FT      /tag= m
FT      /number= 7
FT      exon      7302. .7452
FT      /tag= n
FT      /number= 8
FT      Intron    7453. .7664
FT      /tag= o
FT      /number= 8
FT      exon      7665. .7817
FT      /tag= p
FT      /number= 9
FT      Intron    7818. .7910
FT      /tag= q
FT      /number= 9
FT      exon      7911. .8077
FT      /tag= r
FT      /number= 10
FT      Intron    8078. .9297
FT      /tag= s
FT      /number= 10
FT      exon      9298. .9479
FT      /tag= t
FT      /number= 11
FT      Intron    9480. .10162
FT      /tag= u
FT      /number= 12
FT      exon      10163. .11594
FT      /tag= v
FT      /number= 10
XX
XX      EP787801-A2.
XX
XX      06-AUG-1997.
XX
XX      31-JAN-1997; 97EP-00101622.
XX
XX      01-FEB-1996; 96JP-00016590.
XX
XX      (MITS ) MITSUBISHI CORP.
XX      (MITU ) MITSUBISHI CHEM CORP.
XX
XX      Murase M, Murase J, Hayakawa T, Imamura J, Iwabuchi M;
XX      WPI; 1997-387486/36.
XX
XX      Increasing storage lipid content in seeds and plants - by inhibiting
XX      cytosolic pyruvate kinase.
XX
XX      Claim 9; Page 20-25; 28pp; English.
XX
XX      This DNA encodes a Rattus norvegicus cytosolic pyruvate kinase. The
XX      storage lipid content of a seed is increased by reducing the activity of
XX      endogenous cytosolic pyruvate kinase in the seed. This is applied to
XX      plants which accumulate storage protein and lipid in the embryo,
XX      particularly an oilseed plant such as soya, sunflower, sesame or
XX      CC especially rapeseed. Inhibitory enzymes involved in amino acid
XX      CC biosynthesis increases production of lipid by directing more of the
XX      CC precursor to the chloroplast
XX
XX      Sequence 13011 BP; 3167 A; 3392 C; 3277 G; 3175 T; 0 U; 0 Other;
XX
XX      Query Match      100.0%; Score 51; DB 2; Length 13011;
XX      Best Local Similarity 100.0%; Pred. No. 6.1e-08;
XX      Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      1 CATTGGCGGACGCGGACATCCGCTGCTCTGACTGACCTGCGCCCGACAGTGA 51
XX      ||||||||||||||||||||||||||||||||||||||||||||||||||||
XX      Db      3021 CATGGCGGACGCGGACATCCGCTGCTCTGACTGCGCCCGACAGTGA 3071
XX
XX      RESULT 8
```

```
ABR2448
XX      ID      ABR2448 standard; DNA; 13011 BP.
XX
XX      ABR2448;
XX
XX      26-JUN-2003 (first entry)
XX
XX      Toxicity modelling related rat gene SEQ ID No 2150.
XX
XX      Toxic effect; gene expression profile; renal toxicity; toxicity marker;
XX      KW database; drug screening; toxicity assay; rat; ds.
XX
XX      Rattus norvegicus.
XX
XX      WO200295000-A2.
XX
XX      28-NOV-2002.
XX
XX      22-MAY-2002; 2002MO-US016173.
XX
XX      22-MAY-2001; 2001US-0292335P.
XX      13-JUN-2001; 2001US-0297523P.
XX      19-JUN-2001; 2001US-0298925P.
XX      10-JUL-2001; 2001US-0303807P.
XX      10-JUL-2001; 2001US-0303808P.
XX      10-JUL-2001; 2001US-0303810P.
XX      28-AUG-2001; 2001US-0315047P.
XX      27-SEP-2001; 2001US-0324928P.
XX      22-OCT-2001; 2001US-0330462P.
XX      01-NOV-2001; 2001US-0330867P.
XX      21-NOV-2001; 2001US-0331805P.
XX      06-DEC-2001; 2001US-0336144P.
XX      19-DEC-2001; 2001US-0340873P.
XX      21-FEB-2002; 2002US-0357842P.
XX      21-FEB-2002; 2002US-0357843P.
XX      21-FEB-2002; 2002US-0357844P.
XX      15-MAR-2002; 2002US-0364134P.
XX      08-APR-2002; 2002US-0370144P.
XX      08-APR-2002; 2002US-0370206P.
XX      08-APR-2002; 2002US-0370247P.
XX      17-APR-2002; 2002US-0372794P.
XX      21-APR-2002; 2002US-0371679P.
XX
XX      (GENE-) GENE LOGIC INC.
XX
XX      Wendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX      WPI; 2003-148464/14.
XX
XX      Predicting at least one toxic effect of a compound, useful for toxicity
XX      PT modeling, comprises preparing a gene expression profile of a tissue or
XX      PT cell sample exposed to the compound, and comparing the gene expression
XX      PT profile to a database.
XX
XX      Example 4; Page; 446pp; English.
XX
XX      The invention relates to a novel method of predicting at least one toxic
XX      CC effect of a compound. The method comprises a gene expression profile of a
XX      CC tissue or cell sample exposed to the compound, and comparing the gene
XX      CC expression profile to a database comprising at least part of the data or
XX      CC information given in the specification. The method are useful for
XX      CC predicting at least one toxic effect of a compound, predicting the
XX      CC progression of a toxic effect of a compound, predicting the renal
XX      CC toxicity of a compound, or identifying toxicity markers in tissues or
XX      CC cells exposed to known renal toxin. The genes are useful as toxicity
XX      CC markers in drug screening and toxicity assays, in monitoring disease or
XX      CC physiological states, or disease progression. This polynucleotide
XX      CC represents a rat DNA sequence relating to the toxic effect database
XX      CC described in the specification. NOTE: The sequence data for this patent
XX      CC did not form part of the printed specification, but was obtained in
XX      CC electronic format directly from the World Intellectual Property
XX      CC Organization
```

SQ Sequence 13011 BP; 3167 A; 3392 C; 3277 G; 3175 T; 0 U; 0 Other;
 Query Match 100.0%; Score 51; DB 10; Length 13011;
 Best Local Similarity 100.0%; Pred. No. 6.1e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CATGGGCGACAGGGGACCTCCCGTGTCTCTGAGACTCTGAGCCCCAGTGA 51
 DB 3021 CATGGGCGACAGGGGACCTCCCGTGTCTCTGAGACTCTGAGCCCCAGTGA 3071
 RESULT 9
 ADP72914
 ID ADP72914 standard; DNA; 13011 BP.
 XX
 AC ADP72914;
 XX
 DT 26-AUG-2004 (first entry)
 XX
 DE Renal toxin progression gene marker #1503.
 XX
 KW ds; toxic effect; gene expression profile; kidney tissue;
 KW differential gene expression; toxicity progression; toxicity marker;
 KW drug screening; toxicity assay; kidney pathology; nephritis;
 KW kidney necrosis; glomerular injury; tubular injury;
 KW focal segmental glomerulosclerosis.
 XX
 OS Rattus norvegicus.
 XX
 PN WO2004048598-A2.
 XX
 PD 10-JUN-2004.
 XX
 PF 24-NOV-2003; 2003WO-US037556.
 XX
 PR 22-NOV-2002; 2002US-00301856.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
 PI Elashoff M;
 XX
 DR WPI; 2004-460771/43.
 XX
 PT Predicting (the progression of) a toxic effect of a compound, for
 PT monitoring the progression of renal disease states, comprises preparing a
 PT gene expression profile of a kidney tissue or cell sample exposed to the
 PT compound.
 XX
 PS Claim 11; SEQ ID NO 1503; 266pp; English.
 XX
 CC The invention relates to a method of predicting (the progression of) a
 CC toxic effect of a compound by preparing a gene expression profile of a
 CC kidney tissue or cell sample exposed to the compound and comparing the
 CC gene expression profile to a database, or detecting the level of gene(s)
 CC expression in a tissue or cell sample exposed to the compound, where
 CC differential gene expression compared to a control indicates a toxic
 CC effect (toxicity progression). The method is useful for predicting (the
 CC progression of) at least one toxic effect of a compound. The genes are
 CC useful as toxicity markers in drug screening and toxicity assays. The
 CC methods are useful for predicting the likelihood that a compound or test
 CC agent will induce various specific kidney pathologies, such as nephritis,
 CC kidney necrosis, glomerular and tubular injury, or focal segmental
 CC glomerulosclerosis. The methods are useful for determining the similarity
 CC of a toxic response to one or more individual compounds and for
 CC predicting or elucidating the potential cellular pathways influenced,
 CC induced or modulated by the compound or test agent. The kit is useful for
 CC predicting or modelling the toxic response of a test compound, for
 CC monitoring the progression of renal disease states, for identifying genes
 CC that show promise as new drug targets and for screening known and newly
 CC designed drugs. This sequence corresponds to a gene marker used in the
 CC method of the invention. (Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).
 XX
 SQ Sequence 13011 BP; 3167 A; 3392 C; 3277 G; 3175 T; 0 U; 0 Other;
 Query Match 100.0%; Score 51; DB 12; Length 13011;
 Best Local Similarity 100.0%; Pred. No. 6.1e-08;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CATGGGCGACAGGGGACCTCCCGTGTCTCTGAGACTCTGAGCCCCAGTGA 51
 DB 3021 CATGGGCGACAGGGGACCTCCCGTGTCTCTGAGACTCTGAGCCCCAGTGA 3071
 RESULT 10
 ABX15379
 ID ABX15379 standard; DNA; 48 BP.
 XX
 AC ABX15379;
 XX
 DT 17-APR-2003 (first entry)
 XX
 DE Rat liver pyruvate kinase glucose response element DNA fragment #1.
 XX
 KW Glucose response element; GIRE; liver pyruvate kinase; LPK; ss; IGBP-1;
 KW insulin-sensitive element; ISF; basal promoter; hyperglycaemia; insulin;
 KW insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
 KW fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
 KW glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
 KW hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
 KW intracellular hormone receptor; insulin regulator construct; anabolic.
 XX
 OS Rattus norvegicus.
 XX
 PN US2002107198-A1.
 XX
 PD 08-AUG-2002.
 XX
 PF 10-OCT-2001; 2001US-00972916.
 XX
 PR 11-OCT-2000; 2000US-0239113P.
 XX
 PA (THUL) THULE P M.
 XX
 PI Thule PM;
 XX
 DR WPI; 2002-674190/72.
 XX
 PT New insulin regulator cassette, useful e.g. for treating diabetes,
 PT provides specific, glucose-inducible transgenic expression of insulin in
 PT liver cells.
 XX
 PS Disclosure: Page 10; 37pp; English.
 XX
 CC The invention relates to an insulin regulator construct comprising a
 CC glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
 CC promoter and an insulin-sensitive element (ISE) of an insulin-like growth
 CC factor binding protein-1 (IGBP-1) basal promoter. The construct is used
 CC to treat or prevent diabetic complications, to regulate insulin
 CC production, to modulate hyperglycaemia (without severe hypoglycaemia), to
 CC increase fat catabolism and to retard protein catabolism. The construct
 CC provides stimulation of insulin expression by glucose (but not other
 CC carbohydrates) and glucocorticoids and inhibition by glucagon so that a
 CC combination of these effects can maintain nearly euglycaemic conditions
 CC in diabetics during short-term fasting, large carbohydrate loads or when
 CC fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
 CC thus inhibiting the long-term complications of diabetes. The properties
 CC of the construct are essentially specific for hepatocytes and well-
 CC differentiated hepatoma lines and insulin expression in these cells may
 CC have effects additional to those provided by secreted insulin, e.g.
 CC inhibition of cellular protein degradation, and inhibition, or
 CC stimulation of other intracellular hormone receptors. This sequence
 CC represents a fragment of the rat liver pyruvate kinase glucose response
 CC element, used in production of the construct of the invention

```
XX Sequence 48 BP; 5 A; 17 C; 17 G; 9 T; 0 U; 0 Other;
SQ
Query Match          94.1%; Score 48; DB 6; Length 48;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACAGGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 51
    |||
DB 1 GGGCGACAGGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 48

RESULT 11
ABX15380/c
ID ABX15380 standard; DNA; 51 BP.
XX
AC ABX15380;
XX
XX 17-APR-2003 (first entry)
XX
DE Rat liver pyruvate kinase glucose response element DNA fragment #2.
XX
XX Glucose response element; GIRE; liver pyruvate kinase; LPK; ss; IGFBP-1;
XX insulin-sensitive element; ISE; basal promoter; hyperglycaemia; insulin;
XX insulin-like growth factor binding protein-1; hypoglycaemia; glucose;
XX fat catabolism; protein catabolism; carbohydrate; glucocorticoid; rat;
XX glucagon; euglycaemia; diabetes; fasting; ketogenesis; ketoacidosis;
XX hepatocyte; hepatoma; cellular protein degradation; antidiabetic;
XX intracellular hormone receptor; insulin regulator construct; anabolic.
XX
OS Rattus norvegicus.
XX
XX US2002107198-A1.
XX
PD 08-AUG-2002.
XX
XX 10-OCT-2001; 2001US-00972916.
XX
XX 11-OCT-2000; 2000US-0239113P.
XX
PA (THUL/) THULE P M.
XX
PI Thule PM;
XX
DR WPI; 2002-674190/72.
XX
PT New insulin regulator cassette, useful e.g. for treating diabetes,
PT provides specific, glucose-inducible transgenic expression of insulin in
PT liver cells.
XX
PS Disclosure; Page 10; 37pp; English.
XX
XX The invention relates to an insulin regulator construct comprising a
XX glucose response element (GIRE) of a liver pyruvate kinase (LPK) gene
XX promoter, and an insulin-sensitive element (ISE) of an insulin-like growth
XX factor binding protein-1 (IGFBP-1) basal promoter. The construct is used
XX to treat or prevent diabetic complications, to regulate insulin
XX production, to modulate hyperglycaemia (without severe hypoglycaemia), to
XX increase fat catabolism and to retard protein catabolism. The construct
XX provides stimulation of insulin expression by glucose (but not other
XX carbohydrates) and glucocorticoids and inhibition by glucagon so that a
XX combination of these effects can maintain nearly euglycaemic conditions
XX in diabetics during short-term fasting, large carbohydrate loads or when
XX fed ad libitum and prevent pathological ketogenesis and ketoacidosis,
XX thus inhibiting the long-term complications of diabetes. The properties
XX of the construct are essentially specific for hepatocytes and well-
XX differentiated hepatoma lines and insulin expression in these cells may
XX have effects additional to those provided by secreted insulin, e.g.
XX inhibition of cellular protein degradation, and inhibition, or
XX stimulation of other intracellular hormone receptors. This sequence
XX represents a fragment of the rat liver pyruvate kinase glucose response
XX element, used in production of the construct of the invention
```

```
SQ Sequence 51 BP; 10 A; 17 C; 18 G; 6 T; 0 U; 0 Other;
SQ
Query Match          94.1%; Score 48; DB 6; Length 51;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACAGGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 51
    |||
DB 51 GGGCGACAGGGGCACTCCCGTGTCTCTGACTCTGCCCCCAGTGTGA 4

RESULT 12
ADB83565
ID ADB83565 standard; DNA; 30 BP.
XX
AC ADB83565;
XX
XX 29-JAN-2004 (first entry)
XX
XX ChREBP purification oligonucleotide #2.
XX
XX ss; lipogenesis; glycolysis; glucose metabolism;
XX carbohydrate response element binding protein; ChREBP; obesity; diabetes;
XX vascular disease; carbohydrate metabolism; cardiac disease.
XX
OS Synthetic.
XX
XX US2003124590-A1.
XX
XX 03-JUL-2003.
XX
XX 16-OCT-2002; 2002US-00272206.
XX
XX 16-OCT-2001; 2001US-0329834P.
XX
XX (UYED/) UYEDA K.
XX
PI Uyeda K;
XX
DR WPI; 2004-020312/02.
XX
XX Modulation of the expression of DNA molecule that encodes protein
XX involves use of an agent that induces phosphorylation or
XX PT dephosphorylation of carbohydrate response element binding protein.
XX
XX Example 1; SEQ ID NO 2; 64pp; English.
XX
XX The invention relates to a method of modulation of expression of a DNA
XX molecule (preferably inhibition of lipogenesis or glycolysis) that
XX encodes a protein involved in glucose metabolism or lipogenesis in a cell
XX which comprises contacting the cell with an agent that induces
XX phosphorylation or dephosphorylation of carbohydrate response element
XX binding protein (ChREBP). The method is useful for treating obesity,
XX diabetes or vascular disease and modulating carbohydrate metabolism, for
XX inhibiting lipogenesis or glycolysis and modulating expression of a DNA
XX molecule that encodes a protein involved in glucose metabolism or
XX lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
XX disease. The present sequence represents an oligonucleotide used in the
XX purification of ChREBP
XX
SQ Sequence 30 BP; 3 A; 11 C; 11 G; 5 T; 0 U; 0 Other;
SQ
Query Match          58.8%; Score 30; DB 12; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGGGGCACTCCCGTGTTC 30
    |||
DB 1 CATGGCGCAGGGGCACTCCCGTGTTC 30

RESULT 13
ADB83567/c
```

ID	ADBE83567 standard; DNA; 27 BP.
XX	
AC	ADBE83567;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	L-pyruvate kinase carbohydrate response element #2.
XX	
KM	ds; lipogenesis; glycolysis; glucose metabolism;
KW	carbohydrate response element binding protein; ChREBP; obesity; diabetes;
KM	vascular disease; carbohydrate metabolism; cardiac disease;
XX	L-pyruvate kinase.
OS	Unidentified.
XX	
PN	US2003124590-A1.
XX	
PD	03-JUL-2003.
XX	
PF	16-OCT-2002; 2002US-00272206.
XX	
PR	16-OCT-2001; 2001US-0329834P.
XX	
PA	(UYED/) UYEDA K.
P1	Uyeda K;
XX	
DR	WPI; 2004-020312/02.
XX	
PT	Modulation of the expression of DNA molecule that encodes protein
XX	involves use of an agent that induces phosphorylation or
PT	dephosphorylation of carbohydrate response element binding protein.
XX	
PS	Example 2; SEQ ID NO 4; 64pp; English.
XX	
CC	The invention relates to a method of modulation of expression of a DNA
CC	molecule (preferably inhibition of lipogenesis or glycolysis) that
CC	encodes a protein involved in glucose metabolism or lipogenesis in a cell
CC	which comprises contacting the cell with an agent that induces
CC	phosphorylation or dephosphorylation of carbohydrate response element
CC	binding protein (ChREBP). The method is useful for treating obesity,
CC	diabetes or vascular disease and modulating carbohydrate metabolism, for
CC	inhibiting lipogenesis or glycolysis and modulating expression of a DNA
CC	molecule that encodes a protein involved in glucose metabolism or
CC	lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
CC	disease. The present sequence represents a L-pyruvate kinase carbohydrate
CC	response element.
XX	
SQ	Sequence 27 BP; 4 A; 11 C; 10 G; 2 T; 0 U; 0 Other;
Query Match	52.9%; Score 27; DB 12; Length 27;
Best Local Similarity	100.0%; Pred.No. 7;
Matches 27; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB	
27	GAGGCGCAGGGGCACTCCCGTCGCC 30 GAGGCGCAGGGGCACTCCCGTCGCC 1
RESULT 14	
ADEB83566	
ID	ADEB83566 standard; DNA; 27 BP.
XX	
AC	ADEB83566;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	L-pyruvate kinase carbohydrate response element #1.
XX	
KM	ds; lipogenesis; glycolysis; glucose metabolism;
KW	carbohydrate response element binding protein; ChREBP; obesity; diabetes;
KM	vascular disease; carbohydrate metabolism; cardiac disease;
KW	L-pyruvate kinase.

XX	Unidentified.
OS	
PV	US2003124590-A1.
NN	
PD	03-JUL-2003.
XX	
PF	16-OCT-2002; 2002US-00272206.
XX	
PR	16-OCT-2001; 2001US-0329834P.
XX	
PA	(UYED/) UYEDA K.
XX	
PI	Uyeda K;
XX	
DR	WPI; 2004-020312/02.
XX	
PT	Modulation of the expression of DNA molecule that encodes protein involves use of an agent that induces phosphorylation or dephosphorylation of carbohydrate response element binding protein.
XX	
PS	Example 2; SEQ ID NO 3; 64pp; English.
XX	
CC	The invention relates to a method of modulation of expression of a DNA molecule (preferably inhibition of lipogenesis or glycolysis) that encodes a protein involved in glucose metabolism or lipogenesis in a cell which comprises contacting the cell with an agent that induces phosphorylation or dephosphorylation of carbohydrate response element binding protein (ChREBP). The method is useful for treating obesity, diabetes or vascular disease and modulating carbohydrate metabolism, for inhibiting lipogenesis or glycolysis and modulating expression of a DNA molecule that encodes a protein involved in glucose metabolism or lipogenesis in a cell (e.g. human liver cell) and for treating cardiac disease. The present sequence represents a L-pyruvate kinase carbohydrate response element.
CC	
CC	
CC	
CC	
SQ	Sequence 27 BP; 2 A; 10 C; 11 G; 4 T; 0 U; 0 Other;
XX	
Query Match	52.9%; Score 27; DB 12; Length 27;
Best Local Similarity	100.0%; Pred.No. 7;
Matches	27; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	4 GGCGCAGGGGCACTCCCGTGTTC 30
Db	1 GGCGCAGGGGCACTCCCGTGTTC 27
RESULT 15	
ADE83614	
ID	ADE83614 standard; DNA; 26 BP.
XX	
AC	ADE83614;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	ChREBP gel shift assay gene sequence #1.
XX	
SS	ss; lipogenesis; glycolysis; glucose metabolism; carbohydrate response element binding protein; ChREBP; obesity; diabetes; vascular disease; carbohydrate metabolism; cardiac disease; gel shift assay.
KM	
KW	
XX	
OS	Unidentified.
XX	
PN	US2003124590-A1.
XX	
PD	03-JUL-2003.
XX	
PE	16-OCT-2002; 2002US-00272206.
XX	
PR	16-OCT-2001; 2001US-0329834P.
XX	
PA	(UYED/) UYEDA K.

XX Uyeda K;
 PI
 XX

DR WPI, 2004-020312/02.

XX Modulation of the expression of DNA molecule that encodes protein
 PT involves use of an agent that induces phosphorylation or
 PT dephosphorylation of carbohydrate response element binding protein.

XX Example 14; SEQ ID NO 51; 64pp; English.

PS
 XX The invention relates to a method of modulation of expression of a DNA
 CC molecule (preferably inhibition of lipogenesis or glycolysis) that
 CC encodes a protein involved in glucose metabolism or lipogenesis in a cell
 CC which comprises contacting the cell with an agent that induces
 CC phosphorylation or dephosphorylation of carbohydrate response element
 CC binding protein (ChREBP). The method is useful for treating obesity,
 CC diabetes or vascular disease and modulating carbohydrate metabolism, for
 CC inhibiting lipogenesis or glycolysis and modulating expression of a DNA
 CC molecule that encodes a protein involved in glucose metabolism or
 CC lipogenesis in a cell (e.g. human liver cell) and for treating cardiac
 CC disease. The present sequence represents a ChREBP gel shift assay gene
 CC sequence.
 XX

SO Sequence 26 BP; 2 A; 9 C; 11 G; 4 T; 0 U; 0 Other;

Query Match 51.0%; Score 26; DB 12; Length 26;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGCACTCCCGTGGTTC 29

Db 1 GGGCGACGGGGCACTCCCGTGGTTC 26

Search completed: September 1, 2005, 00:34:35
 Job time : 60.0797 secs

This Page Blank (uspto)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 : Search time 148.534 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-1

Perfect score: 51
Sequence: 1 catggcgccagcgagcgagcgcaccc.....ggagctctgccccagctgta 51

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 733684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	100.0	51	9	US-09-972-916A-1
2	51	100.0	270	9	US-09-972-916A-3
3	51	100.0	321	9	US-09-972-916A-4
4	51	100.0	372	9	US-09-972-916A-5
5	51	100.0	423	9	US-09-972-916A-6
6	51	100.0	13011	17	US-10-388-934-36
7	51	100.0	13011	18	US-10-152-319A-2150

C	8	36.6	71.8	31703	17	US-10-085-117-172	Sequence 172, App
	9	30	58.8	30	15	US-10-272-206-2	Sequence 2, Appli
	10	27	52.9	27	15	US-10-272-206-3	Sequence 3, Appli
	11	27	52.9	27	15	US-10-272-206-4	Sequence 4, Appli
	12	26	51.0	26	15	US-10-272-206-51	Sequence 51, Appli
	13	24.6	48.2	55001	17	US-10-160-497-4	Sequence 4, Appli
	14	24.6	48.2	55001	17	US-10-348-750-4	Sequence 4, Appli
	15	24.6	48.2	55001	21	US-10-991-147-4	Sequence 4, Appli
	16	24.4	47.8	3138	9	US-09-880-107-1716	Sequence 1716, Ap
	17	24.4	47.8	3138	17	US-10-172-118-909	Sequence 909, App
	18	24.4	47.8	3138	18	US-10-342-887-909	Sequence 909, App
	19	24.4	47.8	3138	20	US-10-723-860-176	Sequence 176, App
	20	24.4	47.8	3138	22	US-10-756-149-213	Sequence 213, App
	21	24.4	47.8	3511	20	US-10-723-860-4919	Sequence 4919, App
	22	24	47.1	209	16	US-10-029-386-26227	Sequence 26227, A
	23	24	47.1	525	16	US-10-029-386-12527	Sequence 12527, A
	24	24	47.1	1188	14	US-10-094-417-3	Sequence 3, Appli
	25	24	47.1	1188	24	US-11-086-846-3	Sequence 3, Appli
	26	24	47.1	1191	18	US-10-343-650A-45	Sequence 45, Appli
	27	24	47.1	1311	15	US-10-029-436-1	Sequence 1, Appli
	28	24	47.1	1333	15	US-10-017-161-1921	Sequence 1921, Ap
	29	24	47.1	1498	9	US-09-826-508-37	Sequence 37, Appli
	30	24	47.1	1546	15	US-10-225-567A-584	Sequence 584, App
	31	24	47.1	1579	14	US-10-073-885-11	Sequence 11, Appli
	32	24	47.1	1595	9	US-09-798-710-1	Sequence 1, Appli
	33	24	47.1	1905	21	US-10-505-486-226	Sequence 226, App
	34	24	47.1	2046	9	US-09-764-893-130	Sequence 130, App
	35	24	47.1	2046	9	US-09-764-886-84	Sequence 84, Appli
	36	24	47.1	2046	9	US-09-764-885-375	Sequence 875, App
	37	24	47.1	2046	10	US-09-989-442-155	Sequence 155, App
	38	24	47.1	2046	10	US-09-764-886-84	Sequence 84, Appli
	39	24	47.1	2046	11	US-09-764-885-1209	Sequence 1209, App
	40	24	47.1	2046	14	US-10-073-865-130	Sequence 130, App
	41	24	47.1	2046	14	US-10-103-313-592	Sequence 592, App
	42	23.8	46.7	27	15	US-10-272-206-5	Sequence 5, Appli
	43	23.8	46.7	27	15	US-10-272-206-6	Sequence 6, Appli
	44	23.8	46.7	27	15	US-10-272-206-7	Sequence 7, Appli
	45	23.8	46.7	27	15	US-10-272-206-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-972-916A-1
; Sequence 1, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
US-09-972-916A-1

Query Match 100.0%; Score 51; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 8.1e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCCGTGTCTCGAATCTGGCCCCAGTGA 51
DB 1 CATGGCGCAGCGGCGACTCCCGTGTCTCGAATCTGGCCCCAGTGA 51
RESULT 2

```
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 270;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY
1 CATGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
Db
1 CATGGCGCACGGGCGACCTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
RESULT 3
US-09-972-916A-4/c
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 321;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY
1 CATGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
Db
102 CATGGCGCACGGGCGACCTCCGTTCTCTGACTCTGGCCCCAGTGTGA 52
|||||
```

```
RESULT 4
US-09-972-916A-5/c
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
```

```
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 372;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY
1 CATGGCGCACGGGCACTCCCGTGTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
Db
153 CATGGCGCACGGGCGACCTCCGTTCTCTGACTCTGGCCCCAGTGTGA 103
|||||
```

```
RESULT 5
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
```

```
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 9; Length 423;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY
1 CATGGCGCACGGGCGACCTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
Db
1 CATGGCGCACGGGCGACCTCCGTTCTCTGACTCTGGCCCCAGTGTGA 51
|||||
```

```
RESULT 6
US-10-388-934-36
; Sequence 36, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 13011
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-36
```

```
Query Match
Best Local Similarity 100.0%; Score 51; DB 17; Length 13011;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 CATGGCGCAGCGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 51
Db 3021 CATGGCGCAGCGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 3071

RESULT 7
US-10-152-319A-2150

Sequence 2150, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2150
LENGTH: 13011
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. X05684
US-10-152-319A-2150

Query Match 100.0%; Score 51; DB 18; Length 13011;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 51
Db 3021 CATGGCGCAGCGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 3071

RESULT 8
US-10-085-117-172/C

Sequence 172, Application US/10085117
Publication No. US2003023334A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586

PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 172
LENGTH: 31703
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variation
LOCATION: (1)...(31703)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-172

Query Match 71.8%; Score 36.6; DB 17; Length 31703;
Best Local Similarity 90.7%; Pred. No. 0.00041;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 51
Db 2120 CACGGGCGACTCCGTTGTTCTTGACTCTGCCCCCAGTGA 2078

RESULT 9
US-10-272-206-2

Sequence 2, Application US/102722206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495 0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 30
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-10-272-206-2

Query Match 58.8%; Score 30; DB 15; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCGTTGTTCC 30
Db 1 CATGGCGCAGCGGCGACTCCGTTGTTCC 30

RESULT 10
US-10-272-206-3

Sequence 3, Application US/102722206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495 0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Oligonucleotide from wild type pyruvate kinase
OTHER INFORMATION: carbohydrate response element (CHRE)
US-10-272-206-3

Query Match 52.9%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGGCACTCCCGTGTTC 30
DB 1 GGGCGACGGGGGCACTCCCGTGTTC 27

RESULT 11
US-10-272-206-4/c

Sequence 4, Application US/10272206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495.0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide from wild type pyruvate kinase
OTHER INFORMATION: carbohydrate response element (CHRE)
US-10-272-206-4

Query Match 52.9%; Score 27; DB 15; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGGCACTCCCGTGTTC 30
DB 27 GGGCGACGGGGGCACTCCCGTGTTC 1

RESULT 12
US-10-272-206-51

Sequence 51, Application US/10272206
Publication No. US20030124590A1
GENERAL INFORMATION:
APPLICANT: Kosaku Uyeda
TITLE OF INVENTION: CARBOHYDRATE RESPONSE ELEMENT BINDING
FILE REFERENCE: A34675 090495.0240
CURRENT APPLICATION NUMBER: US/10/272,206
CURRENT FILING DATE: 2002-10-16
PRIOR APPLICATION NUMBER: 60/329,834
NUMBER OF SEQ ID NOS: 66
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 51
LENGTH: 26
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Oligonucleotide
US-10-272-206-51

Query Match 51.0%; Score 26; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGGCACTCCCGTGTTC 29
DB 1 GGGCGACGGGGGCACTCCCGTGTTC 26

RESULT 13
US-10-160-497-4

Sequence 4, Application US/10160497
Publication No. US20030224513A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: ANTISENSE MODULATION OF NOTCH1 EXPRESSION
FILE REFERENCE: RTS-0386
CURRENT APPLICATION NUMBER: US/10/160,497
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 145
SEQ ID NO 4
LENGTH: 55001
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-160-497-4

Query Match 48.2%; Score 24.6; DB 17; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGGCACTCCCGTGTTCGACTGCGCCCACTGT 50
DB 27032 GGGTCACAGGGGCTCTCGCTGCTCCGCCCTCTCCCACTGT 27078

RESULT 14
US-10-348-750-4

Sequence 4, Application US/10348750
Publication No. US20030225019A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Doble
TITLE OF INVENTION: NOTCH1 INHIBITORS FOR INDUCING APOPTOSIS
FILE REFERENCE: ISPH-0729
CURRENT APPLICATION NUMBER: US/10/348,750
CURRENT FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: 10/160,497
NUMBER OF SEQ ID NOS: 146
SEQ ID NO 4
LENGTH: 55001
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-10-348-750-4

Query Match 48.2%; Score 24.6; DB 17; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4 GGGCGACGGGGGCACTCCCGTGTTCGACTGCGCCCACTGT 50
DB 27032 GGGTCACAGGGGCTCTCGCTGCTCCGCCCTCTCCCACTGT 27078

RESULT 15
US-10-991-147-4

Sequence 4, Application US/10991147
Publication No. US20050096292A1
GENERAL INFORMATION:
APPLICANT: Susan M. Freier
APPLICANT: Kenneth W. Doble

```
; APPLICANT: Brich Koller
; TITLE OF INVENTION: NORCH1 INHIBITORS FOR INDUCING APOPTOSIS
; FILE REFERENCE: ISPH-0729
; CURRENT APPLICATION NUMBER: US/10/991,147
; CURRENT FILING DATE: 2004-11-17
; PRIOR APPLICATION NUMBER: US/10/348,750
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 10/160,497
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 146
; SEQ ID NO 4
; LENGTH: 55001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-10-991-147-4
```

```
Query Match          48.2%; Score 24.6; DB 21; Length 55001;
Best Local Similarity 70.2%; Pred. No. 7.5; Mismatches 14; Indels 0; Gaps 0;
Matches 33; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
```

```
Qy      4  GGAGCGACGGGGCACTCCCGTGGTTCCTGGAAGCTGGCCCCAGTGT 50
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      27032 GGATCCACGGGGCTCGCTGGCTTCGGCCCTCTCCCCCACTGTGT 27078
```

```
Search completed: September 1, 2005, 16:44:00
Job time : 164.534 secs
```

This Page Blank (uspto)

Db	943	CCACTGCCCGGAGACACAAACCCAGCAGACTTGACACATGACACAGCGCATCTGCC	1002
QY	181	AGAGAGCTGTGACCACTCCGCTACTAGCTA	214
Db	1003	AGAGAGCTGTGACCACTCCGCTACTAGCTA	1036
RESULT 2			
LOCUS	AC136382		
DEFINITION	AC136382	185148 bp	DNA
ACCESSION	AC136382		
VERSION	AC136382.1	GI:24462257	
KEYWORDS	HTG; HTGS_PHASE1.		
SOURCE	Rattus norvegicus		
ORGANISM	Rattus norvegicus		
REFERENCE	1	(bases 1 to 185148)	
AUTHORS	Murphy, D. Marie, Metzker, M. Lee, Abiraman, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Huilyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokedemeloh, O., Okwundu, G., Olajunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojce, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Soosa, J., Steimle, M., Strong, R., Sutton, A., Sytek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villamasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooten, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.		
TITLE	Unpublished		
JOURNAL	2	(bases 1 to 185148)	
REFERENCE	Rat Genome Sequencing Consortium.		
AUTHORS			

TITLE	Direct Submission
JOURNAL	Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: KDG5 Center clone name: CH230-97018 ----- Summary Statistics Sequencing vector: Plasmid: Chemistry: dye-terminator Big Dye: 100% of reads Assembly program: Phrap; version 0.990329 Consensus quality: 135613 bases at least Q40 Consensus quality: 140849 bases at least Q30 Consensus quality: 145680 bases at least Q20 ----- * NOTE: Estimated insert size may differ from sequence length * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently * consists of 63 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. 1 1010: contig of 1010 bp in length 1011 1110: gap of unknown length 1111 2608: contig of 1498 bp in length 2609 2708: gap of unknown length 2709 3915: contig of 1207 bp in length 3916 4015: gap of unknown length 4016 5547: contig of 1532 bp in length 5548 5647: gap of unknown length 5648 7119: contig of 1472 bp in length 7120 7219: gap of unknown length 7220 8885: contig of 1666 bp in length 8886 10258: gap of unknown length 10259 10358: contig of 1273 bp in length 10359 11626: gap of unknown length 11627 11726: contig of 1268 bp in length 11727 13488: gap of unknown length 13489 15122: contig of 1762 bp in length 15123 15222: gap of unknown length 15223 15946: contig of 1534 bp in length 15947 17046: gap of unknown length 17047 18164: contig of 1118 bp in length 18165 18264: gap of unknown length 18265 19678: contig of 1414 bp in length 19679 19778: gap of unknown length 19779 21687: contig of 1909 bp in length 21688 21787: gap of unknown length 21788 22892: contig of 1105 bp in length 22893 22992: gap of unknown length 22993 24336: contig of 1344 bp in length 24337 24437: gap of unknown length 24438 26617: contig of 2181 bp in length 26618 26717: gap of unknown length 26718 28577: contig of 1860 bp in length 28578 28677: gap of unknown length 28679 30823: contig of 2146 bp in length 30824 30923: gap of unknown length 30924 32266: contig of 1343 bp in length 32267 32366: gap of unknown length 32367 33894: contig of 1528 bp in length 33895 35373: gap of unknown length 35374: contig of 1379 bp in length 35375


```

* 35374 35473: gap of unknown length
* 35474 37295: contig of 1822 bp in length
* 37296 37395: gap of unknown length
* 37396 38420: contig of 1025 bp in length
* 38421 38520: gap of unknown length
* 38521 39974: contig of 1454 bp in length
* 39975 40074: gap of unknown length
* 40075 42283: contig of 2209 bp in length
* 42284 42383: gap of unknown length
* 42384 43721: contig of 1338 bp in length
* 43722 43821: gap of unknown length
* 43822 46538: contig of 2717 bp in length
* 46539 46638: gap of unknown length
* 46639 48621: contig of 1983 bp in length
* 48622 48721: gap of unknown length
* 48722 50602: contig of 1881 bp in length
* 50603 50702: gap of unknown length
* 50703 53841: contig of 3139 bp in length
* 53842 53941: gap of unknown length
* 53942 56195: contig of 2254 bp in length
* 56196 56295: gap of unknown length
* 56296 58949: contig of 2654 bp in length
* 58950 59049: gap of unknown length
* 59050 61032: contig of 1983 bp in length
* 61033 61132: gap of unknown length
* 61133 62942: contig of 1810 bp in length
* 62943 63042: gap of unknown length
* 63043 66045: contig of 3003 bp in length
* 66046 66145: gap of unknown length
* 66146 68197: contig of 2052 bp in length
* 68198 68297: gap of unknown length
* 68298 72235: contig of 3938 bp in length
* 72236 72335: gap of unknown length
* 72336 75803: contig of 3468 bp in length
* 75804 75903: gap of unknown length
* 75904 78355: contig of 2452 bp in length
* 78356 78455: gap of unknown length
* 78456 81458: contig of 3003 bp in length
* 81459 81558: gap of unknown length
* 81559 84673: contig of 3115 bp in length
* 84674 84773: gap of unknown length
* 84774 87739: contig of 2966 bp in length
* 87740 87839: gap of unknown length
* 87840 91535: gap of 3696 bp in length
* 91536 91635: gap of unknown length
* 91636 95522: contig of 3887 bp in length
* 95523 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98830: gap of unknown length
* 98831 101618: contig of 2768 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117631: contig of 3966 bp in length
* 117632 117731: gap of unknown length
* 117732 122502: contig of 4772 bp in length
* 122503 122602: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126645 126744: gap of unknown length

```

Query Match 97.0%; Score 212.4; DB 2; Length 185148;
 Best Local Similarity 99.5%; Pred. No. 2.4e-53;
 Matches 213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTATTATTGTAACACGGGAGATCTAGACACGCTGCTGCACATCA 60
 DB 99253 TCACAGCAAAACAACTATTATTGTAACACGGGAGATCTAGACACGCTGCTGCACATCA 99312
 QY 61 TTAACCGGTGCTGCGAGCCAGCCCTTATATAGGCCCTGGATATGCGCCAGCAGATGCT 120

```

DB 99313 TTAACCGGTGCTGCGAGCCAGCCCTTATATAGGCCCTGGATATGCGCCAGCAGATGCT 99372
QY 121 CCACCTGCCCCGAGACAAACCCAGCAGCATTTGAACTGCACACGCGCATCTGCC 180
DB 99373 CCACCTGCCCCGAGACAAACCCAGCAGCATTTGAACTGCACACGCGCATCTGCC 99432
QY 181 AGAGAGCTGTGACACCACTTCCGCTACTAGCTA 214
DB 99433 AGAGAGCTGTGACACCACTTCCGCTACTAGCTA 99466

RESULT 3
MTLGF
LOCUS MTLGF 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1363)
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
Hepatology 19 (3), 656-665 (1994)
JOURNAL MEDLINE 94164648
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohn,K.L., Waddie,J.R. and Taub,R.
Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
Nucleic Acids Res.
3 (bases 1 to 1363)
REFERENCE Taub,R.A.
AUTHORS Direct Submission
JOURNAL Submitted (23-JUN-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/feature="AP-2 consensus site"
692..706
/feature="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
/feature="TATA_signal"
792..1336
/exon
792
/misc_feature
/note="mRNA cap site"
964..>1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

```

ORIGIN

/crlanslation="MPEFLTVVSWPFLILLSFQIGVAGAPQPMHCACTARLGLCP
PVPASCPBESRPGAGCGCCPTCAIPMGACAGVATPACAGLSGRALPGEPRPLHALTRG
QGASLPEPAPATSTLIFSQHE"

Query Match 86.0%; Score 188.4; DB 10; Length 1363;
Best Local Similarity 92.5%; Pred. No. 4.1e-46;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTTTGAACACGGGATCTTACAGCAGTGGCCCTGACATCA 60
DB 677 TCACAGCAAAACAACTTATTTTGAACACGGGATCTTACAGCAGTGGCCCTGACATCA 736
QY 61 TTAACCCGTGTCGGCAGACACCCCTTCATAGAGCCCTGATGAGCCAGCAGATGAT 120
DB 737 TTAACCTGTGCGGACAGCAGCCCTTCATAGAGCTGTGGTATGACACGACAGATGAT 796
QY 121 CCACGTGCGCGGAGACACAAACCCAGCAGCATTTGAACACTGCACACGCGCATTCGCC 180
DB 797 CCACGTGCGCGGAGACACACACCCAGCAGCATTTGAACACTGCACACGCGCATTCGCC 856
QY 181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 214
DB 857 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 890

RESULT 4
LOCUS AL607124 192843 bp DNA linear ROD 11-APR-2002
DEFINITION Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
ACCESSION AL607124
VERSION AL607124 GI:20145926
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Oliver K.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES

source Location/Qualifiers

1..192843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

ORIGIN

/clone="RP23-20C9"
/clone_11b="RPCI-23"

Query Match 86.0%; Score 188.4; DB 10; Length 192843;
Best Local Similarity 92.5%; Pred. No. 4.5e-46;
Matches 198; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 TCACAGCAAAACAACTTATTTTGAACACGGGATCTTACAGCAGTGGCCCTGACATCA 60
DB 2429 TCACAGCAAAACAACTTATTTTGAACACGGGATCTTACAGCAGTGGCCCTGACATCA 2488
QY 61 TTAACCCGTGTCGGCAGACACCCCTTCATAGAGCCCTGATGAGCCAGCAGATGAT 120
DB 2489 TTAACCTGTGCGGACAGCAGCCCTTCATAGAGCTGTGGTATGACACGACAGATGAT 2548
QY 121 CCACGTGCGCGGAGACACAAACCCAGCAGCATTTGAACACTGCACACGCGCATTCGCC 180
DB 2549 CCACGTGCGCGGAGACACACACCCAGCAGCATTTGAACACTGCACACGCGCATTCGCC 2608
QY 181 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 214
DB 2609 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 2642

RESULT 5
LOCUS BC013345 1569 bp mRNA linear ROD 30-JUN-2004
DEFINITION Mus musculus insulin-like growth factor binding protein 1, mRNA
(cDNA clone MGC:14075 IMAGE:416189), complete cds.
ACCESSION BC013345
VERSION BC013345.1 GI:15426482
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Dere, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dietchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheerz, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrahamson, R.D., Millamy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Guarnatone, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huily, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Heltan, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Merris, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL
PDBMED
REFERENCE
AUTHORS
JOURNAL
REMARK
COMMENT

12477932
2 (bases 1 to 1569)
Strausberg, R.
Direct Submission
Submitted (31-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: http://image.llnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

Location/Qualifiers

1..1569
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:14075 IMAGE:416189"
/cissue_type="Liver, normal. 5 month old male mouse."
/clone_1ib="NCI_CGAP_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocustID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="Insulin-like growth factor binding protein 1"
/protein_id="AAH1345.1"
/db_xref="GI:15426483"
/db_xref="MGI:96436"
/db_xref="LocustID:16006"
/translation="MPEPTLVISWPELILISFOIVAAAGAPQPHWCAPCTARLGLCPYPASPEISRPAGCGCCCTCALPMGACGVAITACAGLSGRLPGEPRPLALTRGQGACVPEPAAPATSTLSSQHEAKAAYASADELSESPMTEQLDSEPHLMAFREDPTLWNAISTVSMRAREIADLKWKEPCORLYKVLLELAQOKAGDELYKFLYLPNCNNGFYHKSQCEITSDGBAGLCWCYPMWGGKIPGSLETRGDPNCHQYFVHN"

CDS

1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocustID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="Insulin-like growth factor binding protein 1"
/protein_id="AAH1345.1"
/db_xref="GI:15426483"
/db_xref="MGI:96436"
/db_xref="LocustID:16006"
/translation="MPEPTLVISWPELILISFOIVAAAGAPQPHWCAPCTARLGLCPYPASPEISRPAGCGCCCTCALPMGACGVAITACAGLSGRLPGEPRPLALTRGQGACVPEPAAPATSTLSSQHEAKAAYASADELSESPMTEQLDSEPHLMAFREDPTLWNAISTVSMRAREIADLKWKEPCORLYKVLLELAQOKAGDELYKFLYLPNCNNGFYHKSQCEITSDGBAGLCWCYPMWGGKIPGSLETRGDPNCHQYFVHN"

ORIGIN

Query Match 44.8%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 9e-19;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 104 TGGCCAGCAGCATGTGTCACGCGCGGAGACACAAACCGAGCATTTGAACACTG 163
DB 20 TGGGCGACGAGCATGTGTCACGCGCGGAGACACACCCAGCATTTGAACACTG 79

QY 164 CACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
DB 80 CACACGCGCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 130

RESULT 6
AX401932 1500 bp DNA linear PAT 06-JUN-2002
LOCUS Sequence 1608 from Patent WO0210453.

DEFINITION AX401932
ACCESSION AX401932
VERSION AX401932.1 GI:21338112

KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Mendrick, D., Porter, M.W., Johnson, K.R., Castle, A.L. and Elashoff, M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;

Gene Logic, Inc. (US)
Location/Qualifiers

FEATURES
source 1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTTCGCTACTAGCTA 214
DB 68 CACTTCGCTACTAGCTA 85

RESULT 7
AX827271 1500 bp DNA linear PAT 12-DEC-2003
LOCUS Sequence 5 from Patent EP1344834.
DEFINITION AX827271
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
F. HOFMANN-LA ROCHE AG (CH)

FEATURES

source
1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN

Query Match 34.9%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 137 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 196
DB 8 CACAAACCCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACAC 67
QY 197 CACTTCGCTACTAGCTA 214
DB 68 CACTTCGCTACTAGCTA 85

RESULT 8

LOCUS RATTGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IgF binding protein-1 (RIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732

KEYWORDS
SOURCE IGF binding protein-1.
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS Mohm, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.

TITLE	The gene encoding rat insulinlike growth factor-binding protein 1 is rapidly and highly induced in regenerating liver
JOURNAL	Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE	91141487
POBMED	1705004
COMMENT	Original
FEATURES	source text: Rat, cDNA to mRNA. location/Qualifiers 1..1500
source	/organism="Rattus norvegicus" /mol_type="mRNA" /strain="Fisher" /db_xref="taxon:10116" /tissue_type="regenerating liver" /dev_stage="adult" 160..978 /codon_start=1 /product="IGF binding protein-1" /protein_id="AAA41380.1" /db_xref="gi:204733" /tranlation="WPERLYTVSWPFLILSFQYRVVAGAPQPMHCACTERLRLCELPVPASCPETISRPPAGCGCCPTCALPIGAACGVAATACAGLSCRALPGSPRLPALTRQGACPAEAPATPSISGSOHEAAVAASDELAESPEMTLEEQLLDSFHMAFSREBPOLIMNAISTYSNMPAREITDLKKWEPCORELVTERLIAAOKKADDELTKFYLPCKNGFHYHSKCQETSIDDEAGLCVMCVWSKKIPGSLFTHTGDPNCHQFNVAQN"
CDS	sig_peptide mac_peptide 160..234 235..975 /product="IGF binding protein-1"
ORIGIN	Query Match 34.9%; Score 76.4; DB 10; Length 1500; Best Local Similarity 98.7%; Pred. No. 3.7e-12; Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	137 CACAACCAGCAGCATTTGAACACTGTGACACGGCCATTCGCCAGAGAGCTGTGACAC 196 DB 8 CACAACCAGCAGCATTTGAACACTGTGACACGGCCATTCGCCAGAGAGCTGTGACAC 67
OY	197 CACTTCCGCCTACTAGCTA 214 DB 68 CACTTCCGCCTACTAGCTA 85
RESULT 9	
AXI63782	
LOCUS	AXI63782 5001 bp DNA linear PAT 22-JUN-2001
DEFINITION	Sequence 46 from Patent WO0138579.
ACCESSION	AXI63782
VERSION	AXI63782.1 GI:14544878
KEYWORDS	
SOURCE	Rattus norvegicus (Norway rat) Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 Gould-Rothberg,B.E., Digipppo,V.A., Ramesh,T.M. and Gerwein,R.W. METHOD OF IDENTIFYING TOXIC AGENTS USING NSAID-INDUCED DIFFERENTIAL GENE EXPRESSION IN LIVER Patent: WO 0138579-A 46 31-MMY-2001; Curagen Corporation (US)
FEATURES	Location/Qualifiers 1..5001 /organism="Rattus norvegicus" /mol_type="unassigned DNA" /db_xref="taxon:10116"
ORIGIN	
Query Match	34.9%; Score 76.4; DB 6; Length 5001;
Best Local Similarity	98.7%; Pred. No. 3.7e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY	137 CACAACCAGCAGCATTTGAACACTGTGACACGGCCATTCGCCAGAGAGCTGTGACAC 196

Db	1	CACAAACCGGAGCATTTGAACATGCACACGGCCATCTGCCAGAGAGCTGTGACCAAC	60
QY	197	CACCTCCGCTACTAGCTA	214
Db	61	CACCTCCGCTACTATCTA	78
RESULT 10			
RATIGFBA			
LOCUS			
DEFINITION	Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,	5001 bp	DNA linear ROD 30-NOV-1995
ACCESSION	complete cds.		
VERSION	L22979.1	GI:1098472	
KEYWORDS	Insulin-like growth factor binding protein-1.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 5001)		
AUTHORS	Lacson, R., Oehler, D., Yang, E., Goswami, R. and Unterman, T.		
TITLE	Dideoxy sequencing and structural analysis of the rat insulin-like		
	growth factor binding protein-1 gene		
JOURNAL	Biochim. Biophys. Acta 1218 (1), 95-98 (1994)		
MEDLINE	94250701		
PUBMED	7514892		
COMMENT	On Nov 30, 1995 this sequence version replaced gi:385167.		
	Original source text: Rattus norvegicus (strain Sprague-Dawley)		
	DNA.		
FEATURES			
source	Location/Qualifiers		
	1..5001		
	/organism="Rattus norvegicus"		
	/mol_type="genomic DNA"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
	/sex="male"		
	/tissue_type="liver"		
	/dev_stage="adult"		
	/tissue_lib="Clontech EMBL-3 SP6/T7"		
	60..525		
exon	/number=1		
5' UTR	60..152		
gene	153..4221		
	/gene="IGFBP-1"		
	join(153..525,1850..2034,2874..3002,4090..4221)		
CDS	/gene="IGFBP-1"		
	/codon_start=1		
	/product="insulin-like growth factor binding protein"		
	/protein_id="AA82581.1"		
	/db_xref="GI:1098473"		
	/translation="MPEFLTVVMPFLILSPVRVAVGAPOPWHCAPTAERLTCLCP		
	PVASQPEISRPAAGCCCPICALGACGVATPAACAGIGLCRALPGEPRLHALTGG		
	QCGACVLEPAPATSSISGSGHEAKAAVAASEDLAESPMTBESQLDSFILMAPSRSD		
	QPLMVAISTYSRAREITDLKRMKPCORELTKVERLAAQAKGADGITYKLYLPN		
	CNNNGVYHSKQCEPSTSDGEAGLCWCYVPMWGKKIPSLLETRGDNCQYENVON"		
	526..1849		
intron	/gene="IGFBP-1"		
	/number=1		
exon	1850..2034		
	/gene="IGFBP-1"		
	/number=2		
intron	2035..2873		
	/gene="IGFBP-1"		
	/number=2		
exon	2874..3002		
	/gene="IGFBP-1"		
	/number=3		
intron	3003..4089		
	/gene="IGFBP-1"		
	/number=3		
exon	4090..4743		
	/number=4		

3' UTR	4222..	.4743	ORIGIN
Query Match	34.9%;	Score 76.4;	DB 10; Length 5001;
Best Local Similarity	98.7%;	Pred. No. 3.7e-12;	
Matches	77;	Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	137	CACAAACCAGGACACTTGAACACTGCACACGCCCATCTGCCCGAGAGCTGTGACAC	196
DB	1	CACAAACCAGGACACTTGAACACTGCACACGCCCATCTGCCCGAGAGCTGTGACAC	60
QY	197	CACCTCCGCTACTACTGA 214	
DB	61	CACCTCCGCTACTACTGA 78	
RESULT 11			
LOCUS	AY560836	1507 bp	mRNA linear ROD 22-MAR-2004
DEFINITION	Spermophilus tridecemlineatus insulin-like growth factor binding protein 1 (IGFBP-1) mRNA, complete cds.		
ACCESSION	AY560836		
KEYWORDS	AY560836.1 GI:45505308		
SOURCE	Spermophilus tridecemlineatus (thirteen-lined ground squirrel)		
ORGANISM	Spermophilus tridecemlineatus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae; Spermophilus.		
AUTHORS	1. (bases 1 to 1507)		
TITLE	Li, Y., Klimants, D. and Hallenbeck, J.M.		
JOURNAL	Cloning and characterization of insulin-like growth factor binding protein 1 (IGFBP-1) from thirteen-lined ground squirrel		
AUTHORS	2. (bases 1 to 1507)		
TITLE	Li, Y., Klimants, D. and Hallenbeck, J.M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (27-FEB-2004) StrobeBranch, NINDS/NIH, 36 Convent Dr., Bethesda, MD 20892, USA		
source	Location/Qualifiers		
gene	1..1507		
CDs	/organism="Spermophilus tridecemlineatus"		
	/mol_type="mRNA"		
	/db_xref="taxon:43179"		
	1..1507		
	/gene="IGFBP1"		
	689..1507		
	/gene="IGFBP1"		
	/codon_start=1		
	/product="insulin-like growth factor binding protein 1"		
	/protein_id="AAS67029.1"		
	/db_xref="GI:45505309"		
	/translation="IMPEYPAAGLWPELILAVOVSTVASSPOMHCAPCSAEKALCP		
	PVPSCEPLSRPAGCGCCCPKALPLGAACGATATRYARGLSCLRLPGSPRFLHALTGG		
	QGACVPEPATPTAGSLSTIEKEKAVSNVPERVPESEKWEETLEQLSFLHMASSSD		
	QPIILNASTYKSRAREMADIKKKKQPCRELYKLERLAKAQKAGEETKYKFLDNP		
	CNKGFFYSKOCETSIDGSEALWCVPWMSGRIIPGSLIETIGDENCHQYENVNQ"		
ORIGIN			
Query Match	34.8%;	Score 76.2;	DB 10; Length 1507;
Best Local Similarity	67.8%;	Pred. No. 4.2e-12;	
Matches	143;	Conservative	0; Mismatches 53; Indels 15; Gaps 2;
QY	2	CACAAGCAAAACAACTATTTTGAACACGGGGATCTCAGCAGCGTGCCTGACAAATAT	61
DB	431	CAGAAGCAAAACAACTATTTTGAACACCTGGGCTCTTGAACACCGCGCTGGCAATAT	490
QY	62	TAACCC-----GTGCTCCGAGCAGACCCCTTCTAAAGCCCTGGGTATGACGAC	111
DB	491	TAACTTTCTAGTCCAAAGTGAGATGAACCGCGCCCTTTTAAAGGCAAGGGCTGACACCGC	550
QY	112	CAGATGTCACATGCCCGCCGAGACACAAACCCAGCAGCATTTGAACACTGCACACGCG	171

Db	551	AAGATGATGTGACTGCCACGGGACATCTCAGAAATTGACACTATCA-----GGCACTGC	605
QY	172	CATTCGCCAGAGAGCTGTGACCACTTC	202
LOCUS			
DEFINITION			
ACCESSION	BC078889	1510 bp	linear
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REMARK			
COMMENT			
FEATURES			
SOURCE			

gene
/clone="MGC:93595 IMAGE:7129185"
/tissue_type="Kidney", rac (Brown Norway)"
/clone_id="NIH_MGC_236"
/lab_host="DH10B"
/note="Vector: pExpress1"
1. .1510
/gene="Igfbp1"
/note="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="LOCUSID:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
142. .960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH78889.1"
/db_xref="GI:50927647"
/db_xref="LOCUSID:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
/translation="MPELTVSWPFLILISQVAVVAGAPQPMHCARCTARLELCP
PVPASCEISRPAAGCCPTCALPLGAACGVATACAGLSCRALPGSRPLHALTRG
OGACVLEPAAPATISSGSHBEAKAAVSEDELASPEMTEBQLDSFHLMAFRED
CPILMNAISTVSMRAREITDLKWKPEPCORELYVLEBLAAQCKADDEIKFYLPN
CMKNGFYHSKQETSIDBAGHCWCYPMGSKKIPGSLERDGPCHQYFNVQN"

ORIGIN

Query Match 29.9%; Score 65.4; DB 10; Length 1510;
Best Local Similarity 98.5%; Pred. No. 7.9e-09;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 148 CGAGATTGACACGACGCGCATCTGCCAGAGACGTGACCACTTCGCGCTA 207
Db 1 CGAGATTGACACGACGCGCATCTGCCAGAGACGTGACCACTTCGCGCTA 60

QY 208 CTAGCTA 214
Db 61 CTATCTA 67

RESULT 13
AY095345 3886 bp DNA linear PRI 30-DEC-2002
LOCUS AY095345
DEFINITION gene, partial cds.
ACCESSION AY095345.1 GI:20853764
VERSION
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
1 (bases 1 to 3886)
Kim,J.J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Utermann,T.G.
Regulation of insulin-like growth factor binding protein-1 promoter
activity by FMR and HOXA10 in primate endometrial cells
Biol. Reprod. 68 (1), 24-30 (2003)
12493691
2 (bases 1 to 3886)
Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
Direct Submission
Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

FEATURES
source 1. .3886
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
<3655. .>3886
/gene="IGFBP-1"

gene

mRNA
<3655. .>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655. .>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAH78889.1"
/db_xref="GI:20853765"
/translation="MSEVPVAVRWLVLLLTVOYGVTVASAPWQAPCSAEKALCPV
PASCSEVTRSGCCPCPCALPLGAACGVATAR"

CDS

ORIGIN

Query Match 28.6%; Score 62.6; DB 9; Length 3886;
Best Local Similarity 64.0%; Pred. No. 5.7e-08;
Matches 135; Conservative 0; Mismatches 59; Indels 17; Gaps 2;

QY 2 CACAAGCAAAACAACTATTATTGACACGCGGATCTCAGACCGCTGCGTGAACATCAT 61
Db 3366 CACTAGCAAAACAACTATTATTGACACCTAGCTCCTAGGTCGCCGCTGCCATCAT 3425

QY 62 TAACCC-----GTGCTGCCGAGCGACCCCTTCATTAAGCCCTGGGTATGCGCAGC 111
Db 3426 TAACCTCCTGGTGAAGTGGCGCGGCTGTGCTTATTAAGGCGCGCTGTGTCCAGC 3485

QY 112 CAGATGTCACATGCGCCCGCGGACACAAACCCAGGAGATTGACACTGCACAGCGC 171
Db 3486 AAGATGCGCCACCGCATTCATTCACGCAAGC-----ATTCGCGCGCGCGCGCGC 3538

QY 172 CATCTGCCAGAGAGCTGTGACCACTTC 202
Db 3539 CACCTCCAGAGACACTGCGCACCGCTCC 3569

RESULT 14
G67139 448 bp DNA linear STS 18-SEP-2000
LOCUS IGFBP1 X1.1 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G67139
ACCESSION G67139.1 GI:10186730
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 448)
Cox,D.G., Bolliot,C. and Canzian,F.
Genome Survey
Unpublished (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Federico Canzian
Genome Analysis Group
International Agency for Research on Cancer
150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France
Tel: +33-4-72738698
Fax: +33-4-72738388
Email: canzian@arc.fr
Primer A: TGCACTGACAAACAAACTT
Primer B: GGGTGACTCCGAGCAGC
STS size: 448
Protocol:
Template: 50 ng
Primer: 4 um each
dNTPs: 2 mM each
MgCl2: 1.5-2.5 mM
Tag: 0.05 units
Total Vol: 25 ul
Buffer: 1.5-2.5 mM
MgCl2: 50 mM
KCl: 20 mM
Tris-HCl: 8.4.
pH:

FEATURES
source 1.448 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind 1.448
primer_bind 1.20
complement(431..448)
ORIGIN

Query Match 27.8%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 1.9e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCTGACATCAT 61
DB 3 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGTGGCGCTGCAATCAT 62
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTATAGGCTTGGGTATGCGCAGC 111
DB 63 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTTTATAGTGGCGCTGTGTCACGC 122
QY 112 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
DB 123 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTGCCCGCGCGCGC 174
QY 171 CCATCTGCCCGAGAGCTGTGACCACTTC 202
DB 175 CCACCTCCCAAGAGCACTGGCCACCGCTCC 206

RESULT 15
LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2394 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:21442452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Alvarez, C., Horne, D., Perez-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-ABR-2002;
GENE LOGIC INC (US)

FEATURES
source Location/Qualifiers
1.6128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN
Query Match 27.8%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 2e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACAGCTGCTGACATCAT 61
DB 473 CACTAGCAAAACAACTTATTTTGAACACTAGCTCTAGCGTGGCGCTGCAATCAT 532
QY 62 TAACCC-----GTGCTGCCGAGCCAGCCCTTATAGGCTTGGGTATGCGCAGC 111
DB 533 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTTTATAGTGGCGCTGTGTCACGC 592
QY 112 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
DB 593 GAGCATCGGCGCACCGCATCC-----CATCCAGCGAGCATCTGCCCGCGCGCGC 644
QY 171 CCATCTGCCCGAGAGCTGTGACCACTTC 202

DB 645 CCACCTCCCAAGAGCACTGGCCACCGCTCC 676

Search completed: September 1, 2005, 03:26:45
Job time : 1371.59 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 1664.98 Seconds
(without alignments)
5066.706 Million cell updates/sec

Title: US-09-972-916B-2

Perfect score: 219
Sequence: 1 tcacagcgaacaaactta.....ttccgctactagtagccgc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.2	44.8	1013	2	BF236974 602026605
2	87.8	40.1	310	6	BY783538 BY783538
3	87.8	40.1	331	6	BY794229 BY794229
4	87.8	40.1	354	6	BY771317 BY771317
5	86.8	39.6	401	1	AT1785818 u1785818
6	86.8	39.6	480	1	AT196314 u1785818
7	86.8	39.6	706	1	AT1530146 u189509.y
8	86.8	39.6	765	1	AT1098594 u189509.y
9	86.8	39.6	785	1	AT1530313 u191401.y
10	86.8	39.6	799	1	AT1529939 u187409.y
11	86.8	39.6	811	6	CA478518 AGENCOURT
12	86.8	39.6	848	1	AT1790802 u1785818
13	86.8	39.6	852	1	AT1528304 u195810.y
14	85.2	38.9	605	1	AT1196154 u169608.y
15	82.2	37.5	380	1	AT1785039 u173406.y
16	80.4	37.2	846	7	CO573026 AGENCOURT
17	80.4	36.7	692	7	CV127049 AGENCOURT
18	79.4	36.3	713	7	CO560662 AGENCOURT
19	75.4	34.4	748	7	CO575629 AGENCOURT
20	73.8	33.7	694	7	CV117001 AGENCOURT
21	73.2	33.4	332	7	W30013 mc24c07.r1
22	72.4	33.1	615	2	BB660958 BB660958
23	72.4	33.1	618	6	CD561711 B0435D02-
24	72.4	33.1	632	1	AT1892189 m188B01.y

25	67.4	30.8	488	1	AA674302 VP96910.r
26	65.4	29.9	801	7	CK472246 AGENCOURT
27	64.4	29.4	840	7	CK473709 AGENCOURT
28	62.8	28.7	759	1	AA105355 mp37d09.r
29	61.6	28.2	269	2	BB604790 BB604790
30	61.6	28.1	587	2	AW916227 EST347531
31	60.8	27.8	1589	3	CR621807 full1-1eng
32	60.8	27.8	1601	3	CR595377 full1-1eng
33	59	26.9	545	1	AA060360 m167a12.r
34	38.8	17.7	498	5	BU684425 UI-CF-ENO
35	36.8	16.8	428	2	BF829782 MR3-HN006
36	36.8	16.8	849	9	AL208895 Telradon
37	35	16.0	342	7	T87256 yd91c06.r1
38	35	16.0	490	5	BX282084 BX282084
39	35	16.0	509	1	AA443181 2x98h10.r
40	35	16.0	532	4	BG283072 602406655
41	35	16.0	533	4	BG388790 602414492
42	35	16.0	561	4	BM772026 K-EST0056
43	35	16.0	602	5	BX371876 BX371876
44	35	16.0	637	4	BM765253 K-EST0046
45	35	16.0	674	1	AA133272 2116h11.r

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:416189 5',
LOCUS BF236974.1 GI:11150891
ACCESSION BF236974.1
VERSION BF236974.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
1 (bases 1 to 1013)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strusberg, Ph.D.
Email: rgapdb-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence at: 581.

FEATURES

source location/Qualifiers
1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:416189"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-SPORT6; Site: 1: NciI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."

ORIGIN

Query Match 44.8%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 1.2e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 104 TGGCAGCAGCAGTCCTCCTGCGCCGCGAGACACCAACCGAGACATTGAACACTG 163

Db	9	TGGGACAGCCAGATGATGTCACCTGCCCCGGGAGACACACCCAGCGAGCATTTGAACACTG	68
Qy	164	CACACGGCCATTTGCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA	214
Db	69	CACACGGCCGCTGTGCCACAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA	119
RESULT 2			
BY783538			
LOCUS			
DEFINITION	BY783538	RIKEN full-length enriched, 17.5 days embryo whole body	EST 23-MAR-2000
ACCESSION	BY783538	Mus musculus cDNA clone U930176D05 5', mRNA sequence.	
VERSION	BY783538		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 310)		
AUTHORS	Carrinzi, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Suganuma, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanishi, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aldonis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakachi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J., and Hayashizaki, Y.		
TITLE	Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289		(2003)
PUBLISHED	22703353		
COMMENT	12819125		
	Contract: Yoshihide Hayashizaki		
	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute		
	The Institute of Physical and Chemical Research (RIKEN)		
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan		
	Tel: 81-45-503-9222		
	Fax: 81-45-503-9216		
	Email: genome-res@sc.riken.jp URL: http://genome.gsc.riken.jp/		
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.		
	Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
	Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES			
source	1. 310		
	Location/Qualifiers		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="U930176D05"		
	/tissue_type="whole body"		
	/dev_stage="17.5 days embryo"		
	/clone_1ib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	40.1%	Score 87.8; DB 6;	Length 310;
Best Local Similarity	92.9%;	Pred. No. 1.3e-15;	
Matches	92;	Conservative 0;	Mismatches 7; Indels 0; Gaps 0;
Qy	116	ATGGCGCACTGGCCCCCGAGACAAACCCACGAGCACTTGAACACTGCACAGCGGCATTC	175
Db	2	ATGGCGCACTGGCCCCCGAGACAAACCCACGAGCACTTGAACACTGCACAGCGGCATTC	61
Qy	176	TGCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA	214

Db	62	TGCCAGAGAGCTGTGACCACTATTGCCACTACTATCTCA	100
RESULT 3	BY794229	331 bp	mRNA
LOCUS	BY794229	RIKEN full-length enriched, 17.5 days embryo whole body	EST 23-MAR-200
DEFINITION	Mus musculus cDNA clone L930292H24 5', mRNA sequence.		
ACCESSION	BY794229		
VERSION	BY794229.1	GI:39720868	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 331)		
AUTHORS	Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Alizawa, K., Arai, K., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watabiki, A., Hirazane, Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kituchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Payan, W., Aitinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakachi, H., Lyons, P., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Hensch, T. K., Brinkmeier, M., Kawaji, J., and Hayashizaki, Y.		
TITLE	Targeting a complex transciptome: the construction of the mouse full-length cDNA encyclopedia		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shuto-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES	Location/Qualifiers		
source	1..331 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="L930292H24" /issue_type="whole body" /dev_stage="17.5 days embryo" /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	40.1%	Score 87.8;	DB 6;
Best Local Similarity	92.9%	Pred. No. 1.3e-15;	
Matches	92;	Conservative 0;	Mismatches 7;
		Indels 0;	Gaps 0;
QY	116	ATGTTCCACTGCGCGCGAGACACAAACCCGAGCATTTGAACACTGCACACGGCCATC	175
DB	2	ATGTGTCACTGCGCGCGAGACACACACCCGAGCATTTGAACACTGCACACGGCCATC	61
QY	176	TGCCAGAGAGCTGTGACCACTATTGCCACTACTATCTCA	214
DB	62	TGCCAGAGAGCTGTGACCACTATTGCCACTACTATCTCA	100
RESULT 4	BY771317		

LOCUS BY771317 354 bp mRNA linear EST 23-MAR-2004
 DEFINITION Mus musculus cDNA clone U930052G15 5', mRNA sequence.
 ACCESSION BY771317
 VERSION BY771317.1 GI:39697955
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Suganara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Wataniki, A., Hirozane-Kitahara, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aldini, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Pagliolini, M., Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayaishizaki, Y.
 TITLE Targeting a complex transcritpome: the construction of the mouse full-length cDNA encyclopedia
 JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)
 MEDLINE 22703353
 PUBMED 12819125
 COMMENT Contact: Yoshihide Hayaishizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9232
 Fax: 81-45-503-9216
 Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.jp/) for further details.

FEATURES
 source
 1..354
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="L930052G15"
 /tissue_type="whole body"
 /dev_stage="17.5 days embryo"
 /clone_1fb="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN
 Query Match 40.1%; Score 87.8; DB 6; Length 354;
 Best Local Similarity 92.9%; Pred. No. 1.4e-15;
 Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 116 ATGGTCACATGCCCGCGGAGACACAACCCAGGAGACATTGAACTGACACAGGCGCATC 175
 |||||
 DB 2 ATGGTCACATGCCCGCGGAGACACAACCCAGGAGACATTGAACTGACACAGGCGCATC 61
 |||||

QY 176 TGCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 214
 |||||
 DB 62 TGCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 100
 |||||

RESULT 5
 A1785818 401 bp mRNA linear EST 02-JUL-1999
 LOCUS A1785818
 DEFINITION u17805.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1888569 5' similar to gb:K15179 M.musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1785818

VERSION A1785818.1 GI:5333534
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Korn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Other ESTs: u17805.x1
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972893
 Seg primer: custom primer used
 High quality sequence stop: 126.

FEATURES
 source
 1..401
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1888569"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1fb="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pMB18-FL3; Site 1: DraIII (CACGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 (ATGCGCCCTTTTCTTTTCTTTT); double-stranded cDNA was ligated to a DraIII adaptor (TCGTCGCTACTG), digested and cloned into distinct DraIII sites of the pMB18-FL3 vector (5' site CACTGTGT, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTTAAGAAGCTGCG and 3' end primer CGACTGCACTCGACGACA."

ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 401;
 Best Local Similarity 92.9%; Pred. No. 2.8e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGATCAGTCCCGCGGAGACACAACCCAGGAGACATTGAACTGACACAGGCGCATC 176
 |||||
 DB 1 TGATCAGTCCCGCGGAGAGACACAACCCAGGAGACATTGAACTGACACAGGCGCATC 60
 |||||

QY 177 GCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 214
 |||||
 DB 61 GCCCAAGAGAGCTGTGACCAACCACTTCGCGTACTACTA 98
 |||||

RESULT 6
 A1196314 480 bp mRNA linear EST 14-OCT-1998
 LOCUS A1196314
 DEFINITION u171a07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1887828 5' similar to gb:M59316 rnal INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA
 sequence.

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lemon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LIND, contact the IMAGE Consortium (info@image.lind.gov) for further information.
 MGI:972152
 FEATURES
 source
 High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI0B"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACTGACGACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 2.9e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGTGTCCACTGCCCCCGAGACACAAACCCAGAGCATTTGAACACTGACACAGGCGCATCT 176
 Db 1 TGTGTCCACTGCCCCCGAGACACACACACCAGACGACATTTGAACACTGACACAGGCGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION u199f09.y1 Sugano mouse liver m1a Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:W5316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146 GI:4444281
 VERSION A1530146.1
 KEYWORDS EST
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LIND, contact the IMAGE Consortium (info@image.lind.gov) for further information.
 MGI:973933
 FEATURES
 source
 High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI0B"
 /note="Organ: liver; Vector: pME18S-Fl3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACTGACGACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 3.1e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 117 TGTGTCCACTGCCCCCGAGACACAAACCCAGAGCATTTGAACACTGACACAGGCGCATCT 176
 Db 1 TGTGTCCACTGCCCCCGAGAGACACACACCAGACGACATTTGAACACTGACACAGGCGCATCT 60
 QY 177 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION u21e1e07.y1 Sugano mouse liver m1a Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
 AUTHORS Maira, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Maira M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930344
 Seq primer: custom primer used
 High quality sequence stop: 395.
 Location/Qualifiers
 1..765
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1481988"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTGCAGCACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 765;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCACTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 176
 Db 1 TGGTCACTGCCCGCGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 60
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 98
 RESULT 9
 LOCUS A1530313 785 bp mRNA linear EST 18-MAR-1999
 DEFINITION u191f01.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1889785 5' similar to gb:M59316.1rai INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
 VERSION A1530313.1 GI:44444448
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
 AUTHORS Maira, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Maira M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence stop: 459.
 Location/Qualifiers
 1..785
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889785"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTGCAGCACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 785;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCACTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 176
 Db 1 TGGTCACTGCCCGCGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 60
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTACTA 98
 RESULT 10
 LOCUS A1529939 799 bp mRNA linear EST 18-MAR-1999
 DEFINITION u167c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1889392 5' similar to gb:M59316.1rai INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI529939
 VERSION AI529939.1 GI:4444074
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 799)
 REFERENCE 1 (bases 1 to 799)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 FEATURES
 source location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pM18-FL3; Site 1: DraIII (CACTGCTG); Site 2: DraIII (CACCAGTGG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TTTGGCTTACTGG], digested and cloned into distinct DraIII sites of the pM18-FL3 vector (5' site CACTGCTG, 3' site CACCAGTGG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTCTAAAGCTGCG and 3' end primer CGACTGCACTGACACACA."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCCACTGCGCCGCGAGACACAAACCAGGAGCATTTGAACACTGCACAGGCCATCT 176
 Db 1 TGGTCCACTGCGCCGCGAGACACACACCAGGAGCATTTGAACACTGCACAGGCCGCTCT 60
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 98
 RESULT 11
 CA478518 811 bp mRNA linear EST 09-MAR-2004
 LOCUS CA478518
 DEFINITION AGENCOURT.10789306 NIH_MGC_152 Mus musculus cDNA clone
 IMAGE:676652 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS
 EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 811)
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs@remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNM00138 row: m column: 23
 High quality sequence stop: 536.
 FEATURES
 source location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCTGCAAGCGCCGCGACACACATTTGACGAAGAAGCGGCTTTTCTTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen. Library amplification in 16 cycles. Library constructed by Mark Bittinger in the Bradfield Laboratory (McCardle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 39.6%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 3.2e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 117 TGGTCCACTGCGCCGCGAGACACAAACCAGGAGCATTTGAACACTGCACAGGCCATCT 176
 Db 39 TGGTCCACTGCGCCGCGAGACACACACCAGGAGCATTTGAACACTGCACAGGCCGCTCT 98
 Oy 177 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 214
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 136
 RESULT 12
 A1790802 848 bp mRNA linear EST 02-JUL-1999
 LOCUS A1790802
 DEFINITION uk28b10.y1 Sugano mouse kidney mla Mus musculus cDNA clone
 IMAGE:1970299 5', similar to gb:281579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 848)
 REFERENCE 1 (bases 1 to 848)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28p10.x1
Contact: Marra M/Waahu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039
Seq primer: custom primer used
High quality sequence stop: 514.

FEATURES

source

1..848
location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse kidney mKia"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CAGCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TTTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CAGCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer CGACTGTGAGCTCGAGCACA."

ORIGIN

Query Match 39.6%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 3.2e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGTTCACATGCCCCCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCACCT 176
|||||
DB 1 TGTTCACATGCCCCCGAGACACACACACCCAGGAGCATTTGAACACTGCACAGCGGCCT 60
|||||
QY 177 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98
|||||

RESULT 13
A1528304 852 bp mRNA linear EST 18-MAR-1999
LOCUS A1528304
DEFINITION u195g10.y1 Sugano mouse liver m1a Mus musculus cDNA clone
IMAGE:1890210.5 similar to gb:W5316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE); mRNA
sequence.

ACCESSION A1528304
VERSION A1528304.1 GI:4442439
KEYWORDS EST.
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Euteleostomi;
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 852)

REFERENCE 1 (bases 1 to 852)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Perion,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/Waahu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534
Seq primer: custom primer used
High quality sequence stop: 478.

FEATURES

source

1..852
location/Qualifiers

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver m1a"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CAGCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TTTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTG, 3' site CAGCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer CGACTGTGAGCTCGAGCACA."

ORIGIN

Query Match 39.6%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 3.2e-15;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 117 TGTTCACATGCCCCCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCACCT 176
|||||
DB 1 TGTTCACATGCCCCCGAGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGGCCT 60
|||||
QY 177 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 214
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98
|||||

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998
LOCUS A1196154
DEFINITION u195d08.y1 Sugano mouse liver m1a Mus musculus cDNA clone
IMAGE:1887665.5 similar to gb:W5316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE); mRNA
sequence.

ACCESSION A1196154
VERSION A1196154.1 GI:3748760
KEYWORDS EST.
ORGANISM Mus musculus (house mouse)
SOURCE Mus musculus
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Euteleostomi;
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 605)

REFERENCE 1 (bases 1 to 605)
Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMNI Mouse EST Project

JOURNAL COMMENT

Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 Washu-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:971987

FEATURES

source

1..605
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887663"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGACACA."

ORIGIN

Query Match

Best Local Similarity 38.9%; Score 85.2; DB 1; Length 605;
 Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 117 TGGTCCACTGCCCGCGAGACACAAACCAGAGAGATTGAACACTGCACAGCGCATCT 176
 |||||
 DB 1 TGGTCCACTGCCCGCGAGACACACACCCAGAGAGATTGAACACTGCACAGCGCATCT 60

QY 177 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 214
 |||||
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98

RESULT 15

A1785039

u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1888018.5' similar to gb:X81579 M.musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION

A1785039
 A1785039.1 GI:5332755

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished (1999)
 The Washu-NCI Mouse EST Project 1999

COMMENT

Other ESTs: u173a06.x1
 Contact: Marra M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972342

FEATURES

source

1..380
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1888018"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGACACA."

ORIGIN

Query Match

Best Local Similarity 37.5%; Score 82.2; DB 1; Length 380;
 Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 120 TCCACTGCCCGCGAGACACAAACCAGAGAGATTGAACACTGCACAGCGCATCTGCCC 179
 |||||

DB 1 TCCACTGCCCGCGAGAGACACACACCCAGAGAGATTGAACACTGCACAGCGCATCTGCCC 60

QY 180 CAGAGAGCTGTGACCCACCACTTCGCTACTAGCTA 214
 |||||

DB 61 CAGAGAGCTGTGACCCACCACTTCGCTACTAGCTA 95

Search completed: September 1, 2005, 06:56:46
 Job time : 1673.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 ; Search time 637.824 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-2

Perfect score: 219

Sequence: 1 tcacagcgaacaaacacta.....ttccgctactagctacgcgc 219

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219	100.0	219	9	US-09-972-916A-2
2	219	100.0	270	9	US-09-972-916A-3
3	219	100.0	321	9	US-09-972-916A-4
4	219	100.0	472	9	US-09-972-916A-5
5	219	100.0	423	9	US-09-972-916A-6
6	76.4	34.9	1500	9	US-09-917-800A-1608
7	76.4	34.9	1500	17	US-10-388-934-5

8	76.4	34.9	1500	17	US-10-191-803-73	Sequence 73, Appl
9	76.4	34.9	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
10	60.8	27.8	6128	9	US-09-880-107-2393	Sequence 2393, Ap
11	60.8	27.8	6128	22	US-10-756-149-1484	Sequence 1484, Ap
12	60.8	27.8	9173	22	US-10-893-315-126	Sequence 126, App
13	60.8	27.8	9174	22	US-10-893-315-160	Sequence 160, App
14	37.8	17.3	7061	15	US-10-311-455-970	Sequence 970, App
15	37.8	17.3	7061	17	US-10-221-613-148	Sequence 148, App
16	35.2	16.1	1687	20	US-10-739-930-5461	Sequence 5461, App
17	35	16.0	769	13	US-10-027-532-164336	Sequence 164336,
18	35	16.0	769	17	US-10-027-532-164336	Sequence 164336,
19	35	16.0	3895	14	US-10-011-585A-76	Sequence 5746, Appl
20	34	15.5	49979	19	US-10-741-601-5746	Sequence 5746, Appl
21	34	15.5	49979	21	US-10-741-601-5746	Sequence 71905, A
22	32.4	14.8	1389	19	US-10-437-963-79182	Sequence 79182, A
23	32.4	14.8	1584	19	US-10-437-963-79185	Sequence 79185, A
24	32	14.6	76180	19	US-10-322-281-492	Sequence 492, Appl
25	31.6	14.4	1129	20	US-10-343-803-31	Sequence 31, Appl
26	31.6	14.4	1288	15	US-10-017-161-2017	Sequence 2017, Ap
27	31.6	14.4	1288	17	US-10-292-798-1663	Sequence 1663, Ap
28	31.6	14.4	1348	17	US-10-120-988-404	Sequence 404, Appl
29	31.6	14.4	1513	18	US-10-072-012-29	Sequence 29, Appl
30	31.6	14.4	117750	19	US-10-367-094-169	Sequence 169, Appl
31	31.4	14.3	360	20	US-10-425-115-28332	Sequence 28332, A
32	31.4	14.3	1134	17	US-10-282-122A-33251	Sequence 33251, A
33	31.2	14.2	668	20	US-10-363-345A-40493	Sequence 40493, A
34	31.2	14.2	668	20	US-10-363-345A-40494	Sequence 40494, A
35	31.2	14.2	668	21	US-10-363-483A-40493	Sequence 40493, A
36	31.2	14.2	668	21	US-10-363-483A-40494	Sequence 40494, A
37	31.2	14.2	955	20	US-10-653-047-2769	Sequence 2769, Ap
38	31.2	14.2	2019	15	US-10-156-761-1314	Sequence 1314, Ap
39	31.2	14.2	6226	17	US-10-221-613-254	Sequence 254, App
40	31.2	14.2	9025608	15	US-10-156-761-1	Sequence 1, Appl
41	31	14.2	466	22	US-10-972-079-82551	Sequence 82551, A
42	31	14.2	2190	21	US-10-696-909A-51	Sequence 51, Appl
43	31	14.2	2247	16	US-10-119-428-34	Sequence 34, Appl
44	31	14.2	2247	17	US-10-291-172-31	Sequence 31, Appl
45	31	14.2	2247	18	US-10-221-278-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-09-972-916A-2
; Sequence 2, Application US/09972916A
; Patent No. US20020107198A1
GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
US-09-972-916A-2

Query Match	100.0%;	Score 219;	DB 9;	Length 219;
Best Local Similarity	100.0%;	Pred. No. 1e-68;		
Matches	219;	Conservative 0;	Mismatches 0;	Indels 0;
		Gaps 0;		
QY	1	TCACAGCAAACTTATTGAAACACGGGATCTTACACAGCTGCTGACATCA	60	
DB	1	TCACAGCAAACTTATTGAAACACGGGATCTTACACAGCTGCTGACATCA	60	
QY	61	TTAACCGGTGTCGAGCAGCCCTCATTAAGCCCTGGGTATGCGACGACATGCT	120	

Db 61 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACGATGCT 120
QY 121 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 180
Db 121 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 180
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219

RESULT 2
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 100.0%; Score 219; DB 9; Length 270;
Best Local Similarity 100.0%; Pred. No. 1.1e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 60
Db 52 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 111
QY 61 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 120
Db 112 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 171
QY 121 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 180
Db 172 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 231
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 232 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 270

RESULT 3
US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 100.0%; Score 219; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.1e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 60
Db 103 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 162
QY 61 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 120
Db 163 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 222
QY 121 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 180
Db 223 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 282
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 283 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 321

RESULT 4
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 100.0%; Score 219; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.2e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 60
Db 154 TCACAGCAAAACAACTTATTTTGAACACGCGGATCTTAGACGCTGCTGACAAATCA 213
QY 61 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 120
Db 214 TTAACCGTGTGCTCCGAGACGACCTTCAATAAGGCCCTGGGTATGCGCACGACATGCT 273
QY 121 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 180
Db 274 CCACGTGCGCGCGAGACACAACCCAGCATTTGAACCTGACACAGCGCATCTGCGCC 333
QY 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 219
Db 334 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGCGCG 372

RESULT 5
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A

```

; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match          100.0%; Score 219; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 1.2e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TCACAAGCAAAACAACTTATTTTGAACGCGGATCTAGACGCTGCGCTACATCA 60
        |||
DB      205 TCACAAAGCAAAACAACTTATTTTGAACGCGGATCTAGACGCTGCGCTACATCA 264
        |||

QY      61 TTAAACCGTGCTGCCGAGCCAGCCCTTCATAGGCGCTGGTATGCGCCAGCCAGCATGT 120
        |||
DB      265 TTAAACCGTGCTGCCGAGCCAGCCCTTCATAGGCGCTGGTATGCGCCAGCCAGCATGT 324
        |||

QY      121 CCACTGCCGCGGAGACAAACCCAGGAGCTTGAACGCTGCAACGCGCATCTGCCCC 180
        |||
DB      325 CCACTGCCGCGGAGACAAACCCAGGAGCTTGAACGCTGCAACGCGCATCTGCCCC 384
        |||

QY      181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGAGCCGC 219
        |||
DB      385 AGAGAGCTGTGACCAACCACTTCGCTACTAGTAGAGCCGC 423
        |||

RESULT 6
US-09-917-800A-1608
; Sequence 1608, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1608
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
```

```

; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match          34.9%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 6.2e-17;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
        |||
DB      8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
        |||

QY      197 CACTCCGCTACTAGCTA 214
        |||
DB      68 CACTCCGCTACTATCTA 85
        |||

RESULT 7
US-10-388-934-5
; Sequence 5, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)
US-10-388-934-5

Query Match          34.9%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 6.2e-17;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      137 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 196
        |||
DB      8 CACAAACCCAGGAGCATTTGAACACTGCAACGCGCATCTGCCAGAGAGCTGTGACAC 67
        |||

QY      197 CACTCCGCTACTAGCTA 214
        |||
DB      68 CACTCCGCTACTATCTA 85
        |||

RESULT 8
US-10-191-803-73
; Sequence 73, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
```

Query Match	34.98;	Score 76.4;	DB 18;	Length 1500;
-------------	--------	-------------	--------	--------------

NUMBER OF SEQ I;

CURRENT APPLICATION NUMBER: US/10/756,149
 CURRENT FILING DATE: 2004-01-12
 NUMBER OF SEQ ID NOS: 5818

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-1484

Query Match      27.8%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      2 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTCAAGCCTGCGCTGACATCAT 61
      1877 CACTAGCAAAACAACTTATTTTGAACACTCAAGCTCTGCGCGCTGCCATCAT 1936
      62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGTATGCCAGC 111
      533 TAACCTCTGTCGTCAGAGTGCGCGGCTGTGCTTTATTAAGTGCGCGCTGTGTCCAGC 592
      112 CAGCATGTGTCACCTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
      593 GAGCATGCGCCACCGCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGCG 644
      171 CCATCTGCCGAGAGCTGTGACCAACCACTTC 202
      645 CCACCTCCCAAGAGACACTGGCCACCGCTCC 676

DB

RESULT 12
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-126

Query Match      27.8%; Score 60.8; DB 22; Length 9173;
Best Local Similarity 66.5%; Pred. No. 4.4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      2 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTCAAGCCTGCGCTGACATCAT 61
      1877 CACTAGCAAAACAACTTATTTTGAACACTCAAGCTCTGCGCGCTGCCATCAT 1936
      62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGTATGCCAGC 111
      1937 TAACCTCTGTCGTCAGAGTGCGCGGCTGTGCTTTTAAGTGCGCGCTGTGTCCAGC 1996
      112 CAGCATGTGTCACCTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
      1997 GAGCATGCGCCACCGCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGCG 2048
      171 CCATCTGCCGAGAGCTGTGACCAACCACTTC 202
      2049 CCACCTCCCAAGAGACTGGCCACCGCTCC 2080

DB

RESULT 13
US-10-893-315-160
```

```

; Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-160

Query Match      27.8%; Score 60.8; DB 22; Length 9174;
Best Local Similarity 66.5%; Pred. No. 4.4e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      2 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTCAAGCCTGCGCTGACATCAT 61
      1877 CACTAGCAAAACAACTTATTTTGAACACTCAAGCTCTGCGCGCTGCCATCAT 1936
      62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCATTAAGCCCTGGGTATGCCAGC 111
      1937 TAACCTCTGTCGTCAGAGTGCGCGGCTGTGCTTTTAAGTGCGCGCTGTGTCCAGC 1996
      112 CAGCATGTGTCACCTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACACGG 170
      1997 GAGCATGCGCCACCGCCATCC-----CATCCAGCAGCATCTGCGCGCGCGCGCGCG 2048
      171 CCATCTGCCGAGAGCTGTGACCAACCACTTC 202
      2049 CCACCTCCCAAGAGACTGGCCACCGCTCC 2080

DB

RESULT 14
US-10-311-455-970/C
; Sequence 970, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 970
; LENGTH: 7061
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-970

Query Match      17.3%; Score 37.8; DB 15; Length 7061;
Best Local Similarity 73.8%; Pred. No. 0.0086;
Matches 48; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
```

```

QY      2 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACGCTGCCCTGACAATCAT 61
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2185 CACTACACAAACAACTTATTTTAAACACTCACTCTTAACGATACGACGCTACCAATCAT 2126
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      62 TAACC 66
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2125 TAACC 2121

```

RESULT 15

```

US-10-221-613-148/C
; Sequence 148, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 148
; LENGTH: 7061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-148

```

```

Query Match      17.3%; Score 37.8; DB 17; Length 7061;
Best Local Similarity 73.8%; Pred. No. 0.0086;
Matches 48; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

```

QY      2 CACAGCAAAACAACTTATTTTGAACACGGGATCTTAGACGCTGCCCTGACAATCAT 61
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2185 CACTACACAAACAACTTATTTTAAACACTCACTCTTAACGATACGACGCTACCAATCAT 2126
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      62 TAACC 66
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2125 TAACC 2121

```

Search completed: September 1, 2005, 16:44:10
 Job time : 647.824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 69.1649 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916B-2

Perfect score: 219
Sequence: 1 tcacacagcaaacacacta.....ttccgctactagctagccgc 219

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.8	27.8	9173	4 US-09-949-001-30	Sequence 30, App1
2	60.8	27.8	9174	4 US-09-949-001-36	Sequence 36, App1
3	35	16.0	601	4 US-09-949-016-19194	Sequence 19194, A
4	35	16.0	601	4 US-09-949-016-73784	Sequence 73784, A
5	35	16.0	2383	4 US-09-949-016-2153	Sequence 2153, App1
6	35	16.0	2408	4 US-09-949-016-74	Sequence 74, App1
7	35	16.0	7561	4 US-09-949-016-11816	Sequence 11816, A
8	35	16.0	7562	4 US-09-949-016-13895	Sequence 13895, A
9	32.2	14.7	4447	2 US-08-304-309-3	Sequence 3, App1
10	32.2	14.7	4447	2 US-08-991-942-3	Sequence 3, App1
11	32	14.6	36180	4 US-09-949-016-11745	Sequence 11745, A
12	32	14.6	36181	4 US-09-949-016-16163	Sequence 16163, A
13	31.6	14.4	1348	4 US-09-949-016-404	Sequence 404, App
14	31.6	14.4	99748	4 US-09-949-016-11990	Sequence 11990, A
15	31.6	14.4	99749	4 US-09-949-016-16518	Sequence 16518, A
16	31.2	14.2	364	4 US-09-270-767-29006	Sequence 29006, A
17	31.2	14.2	601	4 US-09-949-016-168692	Sequence 168692, A
18	31.2	14.2	601	4 US-09-949-016-168693	Sequence 168693, A
19	31.2	14.2	761	4 US-09-270-767-13109	Sequence 13109, A
20	31	14.2	1888	4 US-09-949-016-1341	Sequence 1341, App
21	31	14.2	13145	4 US-09-949-016-13083	Sequence 13083, A
22	30.6	14.0	897	4 US-09-902-540-4899	Sequence 4899, App
23	30.6	14.0	28493	4 US-09-902-540-1241	Sequence 1241, App
24	30.2	13.8	40181	4 US-09-949-016-17016	Sequence 17016, A
25	29.4	13.6	1056	3 US-09-067-782A-1	Sequence 1, App1
26	29.4	13.4	488	4 US-09-621-976-2763	Sequence 2763, App
27	29.2	13.3	1101	4 US-09-205-258-119	Sequence 119, App

c	28	29	13.2	790	3	US-08-998-416-479	Sequence 479, App
c	29	29	13.2	2194	4	US-09-270-767-10411	Sequence 10411, A
c	30	29	13.2	2205	4	US-09-252-991A-11217	Sequence 11217, A
c	31	29	13.2	2454	4	US-09-252-991A-11038	Sequence 11038, A
c	32	29	13.2	13489	4	US-09-949-016-15911	Sequence 15911, A
c	33	29	13.2	20235	1	US-07-642-734C-3	Sequence 3, App1
c	34	29	13.2	20235	3	US-08-439-009A-3	Sequence 3, App1
c	35	28.8	13.2	439	3	US-09-067-782A-3	Sequence 3, App1
c	36	28.8	13.2	777	4	US-09-902-540-3367	Sequence 3367, App
c	37	28.8	13.2	1383	1	US-08-484-04-1	Sequence 1, App1
c	38	28.8	13.2	3182	1	US-08-484-04-11	Sequence 11, App1
c	39	28.8	13.2	3759	1	US-08-252-966B-11	Sequence 11, App1
c	40	28.8	13.2	3785	1	US-08-252-966B-17	Sequence 17, App1
c	41	28.8	13.2	11613	1	US-08-484-04-10	Sequence 10, App1
c	42	28.8	13.2	17125	4	US-09-902-540-1158	Sequence 1158, App
c	43	28.6	13.1	1140	4	US-09-252-991A-901	Sequence 901, App
c	44	28.6	13.1	1615	4	US-09-799-451-116	Sequence 116, App
c	45	28.6	13.1	1765	4	US-09-799-451-117	Sequence 117, App

ALIGNMENTS

```
RESULT 1
US-09-949-001-30
; Sequence 30, Application US/09949001
; Patent No. 6825336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-30

Query Match      27.8%; Score 60.8; DB 4; Length 9173;
Best Local Similarity 66.5%; Pred. No. 8.8e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      2 CACAGCAAAACAACTATTGACACGCGGATCTTACGACCGCTGCCGCAATCAT 61
      1877 CACTGCAAAACAACTATTGACACGCTGCTGCGGCTGCCGCAATCAT 1936
      62 TAACCC-----GTGCTGCCGAGCCAGCCCTTCAAGGCCCTGGATAGCCAGC 111
      1937 TAACCTCTGTGCAAGTGGCGGCGCTGTGCTTTAAGTGGCGGCTGTCCAGC 1996
      112 CAGCATGTTCATCTCCCGCGGAGACACAAACCCAGCAGCATTTGACATTCG-ACACGG 170
      1997 GAGCATCGGCAACCCGCACTCC-----CATTCAGCAGATCTGCGCGCGCGCGCG 2048
QY      171 CCATCGCCAGAGAGCTGTGACCAACCATTC 202
      2049 CCACCTCCAGAGAGCACTGACCAACCATTC 2080

RESULT 2
US-09-949-001-36
; Sequence 36, Application US/09949001
; Patent No. 6825336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
```

FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match 27.8%; Score 60.8; DB 4; Length 9174;
Best Local Similarity 66.5%; Pred. No. 8.8e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 2 CACAGCAAACTTATTTTGAACACGCGGATCTTACAGCTGCTGACAACTAT 61
DB 1877 CACTAGCAAACTTATTTTGAACACTCAGCTCTTACGCGCGCGCTGCAATCAT 1936
QY 62 TAACCC-----GTCTGCCGAGCCAGCCCTTCAATAGCCCTGTGATAGCCAGC 111
DB 1937 TAACCTCTGTGCAAGTGGCGGCTGTGCTTATAGTGGCGCGCTGTCTCAGC 1996
QY 112 CAGCATGTCTCACTGCCCGCGAGACACAACCCAGCAGCATTTGAACACTGC-ACAGCG 170
DB 1997 GAGCATCGGCCACCGCATCC-----CATCAGCAGCATCTGCCCGCGCGCGCG 2048
QY 171 CCATCTGCCAGAGAGCTGTGACACCACTTC 202
DB 2049 CCACCTCCAGAGAGCACTGGCCACGCTCC 2080

RESULT 3
US-09-949-016-19194/c
Sequence 19194, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19194
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-19194

Query Match 16.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 79 CCAGCCCTTCAATAGCCCTGTGATGGCCAGCAGATGCTTCACTGCCCGCCAGAC 138
DB 478 CCAGCCCTTCTGGGGCGCTGTCTCTCCATCCAGCTTAGCCCTCAGAGTTAGGCC 419
QY 139 CAAACCCAGCAGCATTTGAACACTGCACAGCGCCATCTCCAGAGAGCTGTGACCA 198
DB 418 CCATTCAGAGACATCCACCGGATCCAGACCTCTGCACACCACTCATGATGATGCC 359
QY 199 CTT 201
|||

DB 358 CTT 356

RESULT 4
US-09-949-016-73784/c
Sequence 73784, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 73784
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-73784

Query Match 16.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.11;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 79 CCAGCCCTTCAATAGCCCTGTGATGGCCAGCAGATGCTTCACTGCCCGCCAGAC 138
DB 478 CCAGCCCTTCTGGGGCGCTGTCTCTCCATCCAGCTTAGCCCTCAGAGTTAGGCC 419
QY 139 CAAACCCAGCAGCATTTGAACACTGCACAGCGCCATCTGCCAGAGAGCTGTGACCA 198
DB 418 CCATTCAGAGACATCCACCGGATCCAGACCTCTGCACACCACTCATGATGATGCC 359
QY 199 CTT 201
|||
DB 358 CTT 356

RESULT 5
US-09-949-016-2153/c
Sequence 2153, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2153
LENGTH: 2393
TYPE: DNA
ORGANISM: Human
US-09-949-016-2153

Query Match 16.0%; Score 35; DB 4; Length 2393;
Best Local Similarity 55.3%; Pred. No. 0.18;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;


```

; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..3162
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4447
; OTHER INFORMATION: /product= "Pig DPD"
;
; US-08-304-309-3

```

```

Query Match      14.7%; Score 32.2; DB 2; Length 4447;
Best Local Similarity 49.7%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      45 GCTGCCCTGACATCTTAACCCGTCGCGCAGCCGCTTCATTAAGCCCTGGGTAT 104
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2275 GATGGTGTATACGCCACCAACAGCGCTCAGGTCATGAGATTAAAGCCGATGGCAG 2334

QY      105 GGCAGCGCAGCATGTGCCCTGCCCGCGAGACAAACCAGCAGCATTTGAACACTGC 164
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2335 CCTGGCGCAGCGGTGGGTGCTGGCAAGCGGACTATACATAGGAGTGTCTGGCAGCGCC 2394

QY      165 ACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCGTACT 209
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2395 ATCAGACCAATTGCTTTGAGAGCTGTGACCACTTGTCTGTCT 2439

```

```

RESULT 10
; US-08-991-942-3
; Sequence 3, Application US/08991942
; Patent No. 6015673
; GENERAL INFORMATION:
; APPLICANT: GONZALEZ, Frank J.
; APPLICANT: FERNANDEZ-SALGUERO, Pedro
; TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
; TITLE OF INVENTION: DIHYDROXYIMIDINE DEHYDROGENASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza

```

```

; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,942
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..3162
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..4447
; OTHER INFORMATION: /product= "Pig DPD"
;
; US-08-991-942-3

```

```

Query Match      14.7%; Score 32.2; DB 3; Length 4447;
Best Local Similarity 49.7%; Pred. No. 1.9;
Matches 82; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY      45 GCTGCCCTGACATCTTAACCCGTCGCGCAGCCGCTTCATTAAGCCCTGGGTAT 104
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2275 GATGGTGTATACGCCACCAACAGCGCTCAGGTCATGAGATTAAAGCCGATGGCAG 2334

QY      105 GGCAGCGCAGCATGTGCCCTGCCCGCGAGACAAACCAGCAGCATTTGAACACTGC 164
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2335 CCTGGCGCAGCGGTGGGTGCTGGCAAGCGGACTATACATAGGAGTGTCTGGCAGCGCC 2394

QY      165 ACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGCGTACT 209
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2395 ATCAGACCAATTGCTTTGAGAGCTGTGACCACTTGTCTGTCT 2439

```

```

RESULT 11
; US-09-949-016-11745
; Sequence 11745, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

;
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11745
; LENGTH: 36180
; TYPE: DNA
; ORGANISM: Human
; OS-09-949-016-11745

```

Very Match	14.6%	Score 32;	DB 4;	Length 36180;
Local Similarity	53.1%	Pred. No. 5.1;		
Matches	68;	Mismatches	60;	Indels 0;
				Gaps 0;

OY		ACAAATCATTAACCCGCTGCAGACGCCCTTATTAAGGCCCTGGGATATGGCCAGGCCA	113
Db	2652	ACCCCCCTTTCCCCCGCGCGGAAGAACAAGACTGTATGTATGGCCCTTGGGGGCCTTA	2711
OY	114	GCAATGATTCACCTGCCCGCCGAGAACAAACCAGCGACATTGAACACTGCACACGGGCCA	173
Db	2712	GCAGGGTTCACTTCGAGCCAAAGCACATGGCCACTGGCCCCCAGGGAGAAATCCCTTGTTTC	2771
OY	174	TCTGCCCA	181
Db	2772	TCCGCCCA	2779

RESULT 12
US-09-949-016-16163
Sequence 16163, Application US/09949016

```

? APPLICANT VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 16163
? LENGTH: 36181
? TYPE: DNA
? ORGANISM: Human
? OS-09-949-016-16163

```

Query Match	14.6%	Score 32;	DB 4;	Length 36181;
Best Local Similarity	53.1%	Pred. No. 5.1;		
Matches 68;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;

QY	54	ACATTCATATTAACCCGTGTGCGAGGACCGACCCCTTCATTAAGGACCCGTGGTATGGGCAAGCA	113
Db	2652	ACCCCCCTTCCCCCGCGCGGGAAGAACAGAGAGCTAGTAGAGCCCTCGGGTTTGGGGCCCTA	2711
QY	114	GCAATGTCATCTGCGCCCGGAGACACAAACCCAGAGAGACTTTGAACACTGCACAGGCA	173
Db	2712	GCAAGGTTCACTTCGAGGCGAAGCCATGCGCACTGGCCCCAGGGAGAAATCCCTTGTTC	2772
QY	174	TCCTGCCCA	181
Db	2772	TCCTGCCCA	2779

RESULT 13
US-09-774-528-404/C
Sequence 404, Application US/09774528
Patent No. 6740619
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom

```

APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunni
APPLICANT: Dramaac, Radoje T.
TITLE OF INVENTION: No. 6743619e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 802
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: pc_Fl_genes Version 2.0
SEQ ID NO 404
LENGTH: 1348
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (194)..(1285)
US-09-774-528-404

```

Query Match	14.4%	Score 31.6;	DB 4;	Length 1348;
Beet Local Similarity	58.5%	Pred. No. 1.9;		
Matches	55;	Conservative	0;	Mismatches 39;
			Indels	0;
			Gaps	0;

Dy 104 TGCCAGCAGCATGTGCTCACTGGCCCGGAGACACAACACCAGCAGATTGAACACTG 163

Dd 873 TGCCTGCCACCAAGGGCACAGGCCCAACCACTCAGGCCCGAGCCAGTCAGAGCCCCG 814

QY 164 CACACGGCCACTGTGCCAGAGACGTGTGACACC 197
 Db 813 AATACGGCCACGACCGCGGGAGGGGACCCAGC 780

RESULT 14
US-09-949-016-11990/c
; Sequence 11990, Application US/09949016
; Patent No. 6810726

```

? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01107
? CURRENT APPLICATION NUMBER: US/09/949,016
? PRIOR FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,766
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO: 11990
? LENGTH: 99748
? TYPE: DNA
? ORGANISM: Human
? IS-09-949-016-11990

```

Query Match	14.4%	Score 31.6;	DB 4;	Length 99748;
Best Local Similarity	54.2%	Pred. No. 10;		
Matches 64;	Conservative 0;	Mismatches 54;	Indels 0;	Gaps 0;
<p>72 TCCGAGCCAGCCTTCATTAAGGCCCTGGGTATGGCCAGCCAGCATGTCCTGCGCCG 131</p>				

Db 33666 TTCTGAGCAAAACACACAGACCCCAATTGTATACACACCAGAAAGTGAATGATGGG 33607
QY 132 CGAGACACAAACCCAGCCGAGCATTTGAACCTGCACACGSCCATCTGCCAGAGAGCTG 189
Db 33606 CAGGAAACACTCCAGGTAGCTGTGTCTCTCCAGATGACCATTCACCATGAAAAACCG 33549

RESULT 15

US-09-949-016-16518/C
; Sequence 16518, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16518
; LENGTH: 99749
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16518

Query Match 14.4%; Score 31.6; DB 4; Length 99749;

Best Local Similarity 54.2%; Pred.No.10;

Matches 64; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 72 TGCCGAGCCAGCCCTTCATTAAGGCCCTGGGTATGCGCCAGCCAGCATGTGTCACCTGCCGC 131
Db 33666 TTCTGAGCAAAACACACAGACCCCAATTGTATACACACCAGAAAGTGAATGATGGG 33607
QY 132 CGAGACACAAACCCAGCCGAGCATTTGAACCTGCACACGSCCATCTGCCAGAGAGCTG 189
Db 33606 CAGGAAACACTCCAGGTAGCTGTGTCTCTCCAGATGACCATTCACCATGAAAAACCG 33549

Search completed: September 1, 2005, 07:07:07
U00 Time : 73.3315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 13:10:38 ; Search time 1681.14 Seconds

(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-3

Perfect score: 270
Sequence: 1 catggggcagcaggggagctc.....ttccgctactagctagccgc 270

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	212.4	78.7	1181	10	RAT1LGFZ
2	212.4	78.7	185148	2	AC136382
3	188.4	69.8	1363	10	MMILGF
4	188.4	69.8	192843	10	AL607124
5	98.2	36.4	1569	10	BC013345
6	76.6	28.4	1507	10	AY560836
7	76.4	28.3	1500	6	AX401932
8	76.4	28.3	1500	6	AX827271
9	76.4	28.3	1500	10	RAT1LGFZ
10	76.4	28.3	5001	6	AX163782
11	76.4	28.3	5001	10	RAT1LGFZ
12	65.4	24.2	1510	10	BC078889
13	62.6	23.2	3868	9	AY095345
14	60.8	22.5	446	11	GS7139
15	60.8	22.5	6128	6	AX409747
16	60.8	22.5	6128	9	HUM1GFBP1A
17	60.8	22.5	6128	11	G19994
18	60.8	22.5	6480	9	HUM1GFBP1
19	60.8	22.5	9082	9	AY434069

20	60.8	22.5	69887	9	AC091524
21	59.2	21.9	141539	9	AC146152
22	59.2	21.9	189332	2	AC146117
23	59.2	21.9	200935	2	AC148834
24	51	18.9	194	6	AS7715
25	51	18.9	194	6	AR175909
26	51	18.9	13011	6	E14395
27	51	18.9	13011	6	AX827302
28	51	18.9	13011	10	RNLEPG
29	50.4	18.7	895	9	HUM1GFBP1
30	49	18.1	539	11	G67173
31	44.4	16.4	1480	10	RAT1LGFZ
32	43.2	16.0	2717	10	RAT1LGFZ
33	43.2	16.0	231241	2	AC097039
34	37.8	14.0	7061	6	AX251887
35	37.8	14.0	7061	6	AX345899
36	37.8	14.0	7061	6	AX348680
37	36.8	13.6	348997	1	BX640428
38	36.8	13.6	348997	1	BX640427
39	36.8	13.6	349442	1	BX640447
40	36.6	13.6	415	9	HSPYKINPR
41	36.6	13.6	8409	9	HSU47654
42	36.6	13.6	13322	9	AY316591
43	36.6	13.6	13325	9	AL713999
44	35.6	13.2	300217	1	AE016922
45	35	13.0	511	11	BV184757

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus
Rattus norvegicus (Norway rat)
Insulin-like growth factor.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
1 (bases 1 to 1181)
Unterman,T.G., Lacsos,R.G., McGary,E., Whalen,C. and Goswami,R.G.
Biochem. Biophys. Res. Commun. (1991) In press
COMMENT
Original source text: Rattus norvegicus (strain Sprague-Dawley) male adult liver DNA.

FEATURES
source
Location/Qualifiers
1..1181

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

QY	52	TCACAGCAAAACAACTTATTGTAACACGCGGATCTTACAGCGTCCCTGCAATCA	111
DB	823	TCACAGCAAAACAACTTATTGTAACACGCGGATCTTACAGCGTCCCTGCAATCA	882
QY	112	TTAACCGTGTGTCGCGAGCCGCTTATAGGCTTGGTATAGGCCAGCAGCATGCT	171
DB	883	TTAACCGTGTGTCGCGAGCCGCTTATAGGCTTGGTATAGGCCAGCAGCATGCT	942
QY	172	CCATGCGCCGCGAGCAACCAACGAGCATTTGAACATGACACGCGCATCTGCC	231

Db	943	CCACTGCCCCGAGACACAACCCAGGAGCATTTGACACATTCGACAGCGCATCTGCC	1002
Qy	232	AGAGAGCTGTGACCACTCTCCGCTACTAGCTA	265
Db	1003	AGAGAGCTGTGACCACTCTCCGCTACTATCTA	1036
RESULT 2			
AC136382			
LOCUS	AC136382	165148 bp	DNA linear HTG 01-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS		
ACCESSION	AC136382		
VERSION	AC136382.1	GI:24462257	
KEYWORDS	HTG: HTGS_PHASE1		
SOURCE	Rattus norvegicus	(Norway rat)	
ORGANISM	Rattus norvegicus		
REFERENCE	1 (bases 1 to 165148)		
AUTHORS	Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayogaji, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Din, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Faller, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Gava, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idler, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., Kang, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewe, L., Louisedge, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojce, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherter, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejor, Z., Uemari, K., Valse, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlaczek, R., Wooten, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 165148)		
AUTHORS	Rat Genome Sequencing Consortium.		

TITLE	Direct Submission
JOURNAL	Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu Center project name: KDGS Center clone name: CH230-97018 ----- Summary Statistics Sequencing vector: Plasmid: Chemistry: Dye-terminator Big Dye: 100% of reads Assembly program: Phrap, version 0.990329 Consensus quality: 135613 bases at least Q40 Consensus quality: 140849 bases at least Q30 Consensus quality: 145680 bases at least Q20 -----
	* NOTE: Estimated insert size may differ from sequence length * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html). * NOTE: This is a 'working draft' sequence. It currently * consists of 63 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.
1	1010: contig of 1010 bp in length
1011	1110: gap of unknown length
1111	2608: contig of 1498 bp in length
2609	2708: gap of unknown length
2709	3915: contig of 1207 bp in length
3916	4015: gap of unknown length
4016	5547: contig of 1532 bp in length
5548	5647: gap of unknown length
5649	7119: contig of 1472 bp in length
7120	7219: gap of unknown length
7220	8885: contig of 1666 bp in length
8886	8985: gap of unknown length
8987	10258: contig of 1273 bp in length
10259	10358: gap of unknown length
10359	11626: contig of 1268 bp in length
11627	11726: gap of unknown length
11727	13488: contig of 1762 bp in length
13489	13588: gap of unknown length
13589	15122: contig of 1534 bp in length
15123	15222: gap of unknown length
15223	15946: contig of 1724 bp in length
15947	17046: gap of unknown length
17047	18164: contig of 1118 bp in length
18165	18264: gap of unknown length
18265	19678: contig of 1414 bp in length
19679	19778: gap of unknown length
19779	21687: contig of 1909 bp in length
21688	21787: gap of unknown length
21788	22892: contig of 1105 bp in length
22893	22992: gap of unknown length
22993	24336: contig of 1344 bp in length
24337	24436: gap of unknown length
24437	26517: contig of 2181 bp in length
26518	26717: gap of unknown length
26719	28577: contig of 1860 bp in length
28578	28677: gap of unknown length
28679	30823: contig of 2146 bp in length
30824	30923: gap of unknown length
30924	32265: contig of 1343 bp in length
32266	32366: gap of unknown length
32367	33894: contig of 1528 bp in length
33895	33994: gap of unknown length
33995	35373: contig of 1379 bp in length

```
* 35374 35473: gap of unknown length
* 35474 37295: contig of 1822 bp in length
* 37296 37395: gap of unknown length
* 37396 38420: contig of 1025 bp in length
* 38421 38520: gap of unknown length
* 38521 39974: contig of 1454 bp in length
* 39975 40074: gap of unknown length
* 40075 42283: contig of 2209 bp in length
* 42284 42383: gap of unknown length
* 42384 43721: contig of 1338 bp in length
* 43722 43822: gap of unknown length
* 43822 46538: contig of 2717 bp in length
* 46539 46638: gap of unknown length
* 46639 48621: contig of 1983 bp in length
* 48622 48721: gap of unknown length
* 48722 50602: contig of 1881 bp in length
* 50603 50702: gap of unknown length
* 50703 53841: contig of 3139 bp in length
* 53842 53941: gap of unknown length
* 53942 56195: contig of 2254 bp in length
* 56196 56295: gap of unknown length
* 56296 58949: contig of 2654 bp in length
* 58950 61032: contig of 1983 bp in length
* 61033 61132: gap of unknown length
* 61133 62942: contig of 1810 bp in length
* 62943 63042: gap of unknown length
* 63043 66045: contig of 3003 bp in length
* 66046 66145: gap of unknown length
* 66146 68197: contig of 2052 bp in length
* 68198 68298: gap of unknown length
* 68299 72235: contig of 3938 bp in length
* 72236 72335: gap of unknown length
* 72336 75803: contig of 3468 bp in length
* 75804 75904: gap of unknown length
* 75905 78355: contig of 2452 bp in length
* 78356 78455: gap of unknown length
* 78456 81458: contig of 3003 bp in length
* 81459 81558: gap of unknown length
* 81559 84673: contig of 3115 bp in length
* 84674 84773: gap of unknown length
* 84774 87739: contig of 2966 bp in length
* 87740 87839: gap of unknown length
* 87840 91535: contig of 3636 bp in length
* 91536 91635: gap of unknown length
* 91636 95622: contig of 3887 bp in length
* 95623 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98830: gap of unknown length
* 98831 101618: contig of 2768 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117630: contig of 3966 bp in length
* 117631 117730: gap of unknown length
* 117731 122502: contig of 4772 bp in length
* 122503 122602: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126645 126744: gap of unknown length
```

Query Match 78.7%; Score 212.4; DB 2; Length 185148;
Best Local Similarity 99.5%; Pred. No. 3e-52; Indels 0; Gaps 0;
Matches 213; Conservative 0; Mismatches 1;

QY 52 TCACAGCAAAACAACTTATTTTGAACACGGGGATCTAGACGCTGCTGACAAATCA 111
|TCACAGCAAAACAACTTATTTTGAACACGGGGATCTAGACGCTGCTGACAAATCA 111
DB 99253 TCACAGCAAAACAACTTATTTTGAACACGGGGATCTAGACGCTGCTGACAAATCA 99312
QY 112 TTAACCGCTGCTGCGAGCCAGCCCTTCTAATAGGCGCTGCGATATGCGCAGCATGCT 171

```
DB 99313 TTAACCGCTGCTGCGAGCCAGCCCTTCTAATAGGCGCTGCGATATGCGCAGCATGCT 99372
QY 172 CCACCTGCGCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATCTGCC 231
DB 99373 CCACCTGCGCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACACGCGCATCTGCC 99432
QY 232 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 265
DB 99433 AGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 99466

RESULT 3
MILIF
LOCUS 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1363)
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression
JOURNAL Hepatology 19 (3), 656-665 (1994)
MEDLINE 94164648
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohn,K.L., Waddell,J.R. and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
JOURNAL Nucleic Acids Res.
REFERENCE 3 (bases 1 to 1363)
AUTHORS Taub,R.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curtie Boulevard, Philadelphia, PA 19104-6145, USA

FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
exon
792..1336
/number=1
792
/note="mRNA cap site"
964..>1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"
```


DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

contact: amadan@systemsbio.org
Anup Madan, Jessica Fahy, Erin Helton, Mark Keltman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN.ac: http://image.lnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1.1569

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:14075 IMAGE:4161889"

/issue_type="Liver, normal. 5 month old male mouse."

/clone_lib="NCI CGAP L19"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene

1.1569

/gene="Igfbp1"

/note="Synonym: IGFBP-1"

/db_xref="LocustID:16006"

/db_xref="MGI:96436"

204.1022

/gene="Igfbp1"

/codon_start=1

/product="Insulin-like growth factor binding protein 1"

/protein_id="AAH1345.1"

/db_xref="GI:15426483"

/db_xref="MGI:16006"

/db_xref="MGI:96436"

/translation="MPELVTVSWPELILSFOIGVAAAGAPQPMHCAIPCTARLGLCP

PVPASCPBLSRPAAGCCCTCTCALPMGACGATATACAGLSCRALPGRRPLHALTRG

QACVPEPAATSTLSSQHEEAKAAYVADLSBSPMTSEQLDSEHLMASSED

OPTLMAISTYSMRAREIADIKWKEPCRELKYVLERLAAQKAGBEIKFYLPN

CMKNGFYHSKQCEISLDGEALCWCYPMWGRIPGSLERDPPNCHQYFNVN"

ORIGIN

Query Match 36.4%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 3.2e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 155 TGGCCAGCAGATGCTCCAGTCCCGCGAGACACAAACCGAGAGATTGAACACTG 214
|||
DB 20 TGGGCGAGCAGATGCTCCAGTCCCGCGAGACACACCCGCGAGATTGAACACTG 79
|||
QY 215 CACACGGCCATCTGCCAGAGAGCTGTGACCACTTCGGTACTAGCTA 265
|||
DB 80 CACAGGGCGCTCTGCCAGAGAGCTGTGACCACTTCGACTACTATCTA 130
|||

RESULT 6 1507 bp mRNA linear ROD 22-MAR-2004
LOCUS AY560836
DEFINITION Spermophilus tridecemlineatus insulin-like growth factor binding
protein 1 (IGFBP1) mRNA, complete cds.
ACCESSION AY560836
VERSION AY560836.1 GI:45505308
KEYWORDS
SOURCE Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
ORGANISM Spermophilus tridecemlineatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuriidae; Scuriinae;
Spermophilus.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Li,Y., Klimanis,D. and Hallenbeck,J.M.
TITLE Cloning and characterization of insulin-like growth factor binding
protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1507)
AUTHORS Li,Y., Klimanis,D. and Hallenbeck,J.M.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
Bethesda, MD 20892, USA

FEATURES

source

1.1507

/organism="Spermophilus tridecemlineatus"

/mol_type="mRNA"

/db_xref="taxon:43179"

1.1507

/gene="IGFBP1"

689.1507

/gene="IGFBP1"

/codon_start=1

/product="Insulin-like growth factor binding protein 1"

/protein_id="AA567029.1"

/db_xref="GI:45505309"

/translation="MPEVPAAGLMPFLLLAIVQSTVASSTQPMHCAIPCAKALACP

FVPSCPBLSRPAAGCCCPMCALPGAAGVATAYAGLSCRALPGRRPLHALTRG

QACVPEPAATSTLSSQHEEAKAAYVADLSBSPMTSEQLDSEHLMASSED

OPTLMAISTYSMRAREIADIKWKEPCRELKYVLERLAAQKAGBEIKFYLPN

CMKNGFYHSKQCEISLDGEALCWCYPMWGRIPGSLERDPPNCHQYFNVN"

ORIGIN

Query Match 28.4%; Score 76.6; DB 10; Length 1507;
Best Local Similarity 67.6%; Pred. No. 8.8e-12;
Matches 144; Conservative 0; Mismatches 54; Indels 15; Gaps 2;

QY 51 ATCAGAGCAAAACAACTTATTTTGAACAGGGATCTAGCAGCGCCCTGCAATC 110
|||
DB 429 AGCAGAGCAAAACAACTTATTTTGAACAGGGATCTAGCAGCGCCCTGCAATC 488
|||
QY 111 ATTA-----CCGCTGCTGCGAGCCAGCCCTCATPAGSCCTGGGTATGGCA 160
|||
DB 489 ATTACTTTTCACTGCTGAGAGGAAACCGCCCTTATTAAGCAAGGCTGCACCA 548
|||
QY 161 GCCAGCATGCTCACTGCGCGCGAGACAAACCCAGCAGATTGAACACTGCACG 220
|||
DB 549 GCAAGCATGCTGCACTGCGCGAGCAGGAGCTCAGAAATTGACACTATCA-----GCCACT 603
|||
QY 221 GCCATGCTGCCAGAGAGCTGTGACCAACCACTTC 253
|||
DB 604 GCCAGCTGCCAGAGAGATTGAACCACTGTCC 636
|||

RESULT 7 1500 bp DNA linear PAT 06-JUN-2002
LOCUS AX401932
DEFINITION Sequence 1608 from Patent WO0210453.
ACCESSION AX401932
VERSION AX401932.1 GI:21338112
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1
AUTHORS Mendrick,D., Porter,M.W., Johnson,K.R., Caetle,A.L. and
Elashoff,M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;
Gene Logic, Inc. (US)

FEATURES

Location/Qualifiers

1.1500

source

/organism="Rattus norvegicus"

ORIGIN

/mol_type="unassigned DNA"

TITLE

/db_xref="taxon:10116"

REFERENCE

/note="EMBL/GenBank Accession No. NM_013144"

Query Match	Similarity	28.3%;	Score	76.4;	DB	6;	Length	1500;	
Best Local	Similarity	98.7%;	Pred. No.	1e-11;					
Matches	77;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	188	CACAAACCCAGGACGACATTGAACACTGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	247						
Db	8	CACAAACCCAGGACGACATTGAACACTGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	67						
Qy	248	CACCTTCGGCTACTACTA	265						
Db	68	CACCTTCGGCTACTACTA	85						
RESULT 8									
AX827271									
LOCUS	AX827271		1500 bp	DNA	linear	PAT 12-DEC-2003			
DEFINITION	Sequence 5 from Patent EP1344834.								
ACCESSION	AX827271								
VERSION	AX827271.1		GI:39837360						
KEYWORDS									
SOURCE									
ORGANISM	Rattus norvegicus (Norway rat)								
REFERENCE									
AUTHORS	Boess,F., Suter-Dick,L. and Wolf,D.								
TITLE	Methods for the toxicity prediction of a compound								
JOURNAL	Patent: EP 1344834-A 5 17-SEP-2003;								
FEATURES	F. HOFMANN-LA ROCHE AG (CH)								
source	location/Qualifiers								
1..1500									
/organism="Rattus norvegicus"									
/mol_type="unassigned DNA"									
/db_xref="taxon:10116"									
ORIGIN									
Query Match	28.3%;	Score	76.4;	DB	6;	Length	1500;		
Best Local	Similarity	98.7%;	Pred. No.	1e-11;					
Matches	77;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	188	CACAAACCCAGGACGACATTGAACACTGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	247						
Db	8	CACAAACCCAGGACGACATTGAACACTGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	67						
Qy	248	CACCTTCGGCTACTACTA	265						
Db	68	CACCTTCGGCTACTACTA	85						
RESULT 9									
RAT1GFB									
LOCUS	RAT1GFB		1500 bp	mRNA	linear	ROD 27-APR-1993			
DEFINITION	Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.								
ACCESSION	M58634								
VERSION	M58634.1		GI:204732						
KEYWORDS	IGF binding protein-1.								
SOURCE	Rattus norvegicus (Norway rat)								
ORGANISM	Rattus norvegicus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;								
TITLE	Rattus.								
JOURNAL	1 (bases 1 to 1500)								
MEDLINE	Mohn,K.L., Melby,A.E., Tewari,D.S., Laz,T.M. and Taub,R.								
PUBMED	The gene encoding rat insulinlike growth factor-binding protein 1								
COMMENT	is rapidly and highly induced in regenerating liver								
FEATURES	Mol. Cell. Biol. 11 (3), 1393-1401 (1991)								
source	1705004								
location/Qualifiers	Original								
1..1500									
/organism="Rattus norvegicus"									

		/mol_type="rRNA"
		/strain="fisher"
		/db_xref="taxon:10116"
		/tissue_type="regenerating liver"
		/dev_stage="adult"
		160..978
		/codon_start=1
		/product="IGF binding protein-1"
		/protein_id="AA41380.1"
		/db_xref="GI:204733"
		/translation="MPEFLTVSWPFIILSFQYRVVAGAPQMHCAPCTAERLTCLF VPSPASPIRPGCCGCPICALPLGAACGYATRCAGGSCRALPGEPRLHLTRG OGAVLPAAPATSSLSGSOEAKAIVASDELAESEPMTEEQLISFHLMAAREI OPTLMALISTSMRARBITDCLKKKEPCQRELYKVLERLLAAQQKADELYKYLYLR CNKKGFPHSKQETSLLDEGLCWCVTVPMSKKLPGLSLTRKDENCHQYFNVN"
		160..234
		sig_peptide
		mat_peptide
		235..975
		/product="IGF binding protein-1"
ORIGIN		
Query Match	28.3%; Score 76.4; DB 10; Length 1500;	
Best Local Similarity	98.7%; Pred. No. 1e-11;	
Matches	77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
Oy	188 CACAACCAGCGAGCATTTGAACACTGCACACGGCCATTGTGCCAGAGAGCTGTGACCAAC	247
Db	1 CACAAACCCAGCGAGCATTTGAACACTGCACACGGCCATTGTGCCAGAGAGCTGTGACCAAC	60
Oy	248 CACTTCGGCTACTACTA	265
Db	61 CACTTCGGCTACTACTA	78
RESULT 11		
RATIGFBA		

Db	1	CACAAACCCAGCGAGCATTTGAACACTGCACAGCGCCATCTGCCCAGAGCTGTGACAC	60
Qy	248	CACCTCGCTACTAGCTA	265
Db	61	CACCTCCGCTACTATCTTA	78
RESULT 12			
LOCUS	BC078889		
DEFINITION	BC078889	1510 bp	mRNA linear ROD 03-AUG-2004
	Rattus norvegicus insulin-like growth factor binding protein 1,		
	mRNA (cDNA clone MGC:93595 IMAGE:7129185), complete cds.		
ACCESSION	BC078889		
VERSION	BC078889.1	GI:50927646	
KEYWORDS	MGC.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
DIFFERENCE	1	(bases 1 to 1510)	

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REMARK	
1 (bases 1 to 1510)	Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Krauser, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.R., Muzny, D.M., Sodergren, E., J., Lu, X., Gibbs, R.A., Rahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalsky, J., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	12477932	2 (bases 1 to 1510) 2 Director MGC Project. Direct Submission	Submitted (02-AUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	NIH-MGC Project URL: http://mgc.ncl.nih.gov Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Howard Jacobs cDNA Library Preparation: Express Genomics cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing I.M.A.G.E. at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-hgsc.stanford.edu Contact: (Dickson, Mark) med@axl1.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 184 Rowe Column: 23

```
FEATURES
source
passed the following selection criteria: matched mRNA gi: 69981079.
Location/Qualifiers
1..1510
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
```

gene

clone="MGC:93595 IMAGE:7129185"
/cissue_type="Kidney, rat (Brown Norway)"
/clone_lib="NIH_MGC_236"
/lab_host="DH10B"
/note="Vector: pExpress1"
1. .1510
/gene="Igfbp1"
/note="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="locusID:25685"
/db_xref="RATMAP:44422"
/db_xref="RBD:2872"
142. .960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH78889.1"
/db_xref="GI:50927647"
/db_xref="locusID:25685"
/db_xref="RATMAP:44422"
/db_xref="RBD:2872"
/translation="MPEFLTVSWPFLILSPQVRVAVAGAPQPHCAPCTAERLICP
PVPASCPISRRAGGCGCTCALPIGAACGATATACAGLSCRALPGRPRHALTRG
GACVLEPAAPATSSLSQHEAPAAVAEDLESPBMTBQULDSFHLMAPRED
OPTLMAISTYSMRAEITDLKWKKECORELYVLERLAAQKADDELYKFLPN
CNKNGFVHSKQETSLDGEAGLCWCVFWSGKILPGSLETRGDPNCHQYFNVON"

ORIGIN

Query Match 24.2%; Score 65.4; DB 10; Length 1510;
Best Local Similarity 98.5%; Pred. No. 1.9e-08;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 CGAGATTGAACACGACACGCGCCATCTGCGCCAGAGAGCTGTGACCACTTCCGCTA 258
|||||
1 CGAGATTGAACACGACACGCGCCATCTGCCCAAGAGCTGTGACCACTTCCGCTA 60

QY 259 CTAGCTA 265
|||||
61 CTATCTA 67

Db

RESULT 13
AY095345 3886 bp DNA linear PRI 30-DEC-2002
LOCUS Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
DEFINITION
ACCESSION AY095345
VERSION AY095345.1 GI:20853764
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
1 (bases 1 to 3886)
Kim,J.J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
Regulation of insulin-like growth factor binding protein-1 promoter
activity by FKHR and HOKA10 in primate endometrial cells
Biol. Reprod. 68 (1), 24-30 (2003)
JOURNAL
PUBMED 12493691
2 (bases 1 to 3886)
Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
Direct Submision
Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

FEATURES
source 1. .3886
location/Qualifiers
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
<3655. .>3886
/gene="IGFBP-1"

gene

mRNA

<3655. .>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655. .>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAH23273.1"
/db_xref="GI:20853765"
/translation="MSEVPARVWLVLITVQGVATASAPWQCAPSPCAERLALCPV
PASCSEVTRSHAGCGCCPMCALPIGAAGVATAR"

CDS

ORIGIN

Query Match 23.2%; Score 62.6; DB 9; Length 3886;
Best Local Similarity 64.0%; Pred. No. 1.3e-07;
Matches 135; Conservative 0; Mismatches 59; Indels 17; Gaps 2;

QY 53 CACAGCAAAACAATTATTTTGAACACGCGGATCTTAGACGCTGCGCTGACATCAT 112
|||||
3366 CACTAGCAAAACAATTATTTTGAACATCAGCTCCTTAGCGTGCGCGCTGCCAATCAT 3425

QY 113 TAAACC-----GTGTCGCCGAGCCAGCCCTTATAGGCGCTGGGATAGCCAGC 162
|||||
3426 TAACTCTGTGTGCAAGTGGCGCGCGCTGTGCTTTATAGGCGCGCTGTGTCCAGC 3485

QY 163 CAGCATGGTCCACATGCGCGCGCGAGACAACAACCGAGAGATTGAACATGCACACGGC 222
|||||
3486 AAGCATGGCGCCACCGCCATCTTCATCCAGCANAGC-----ATTCGCCGCGCGCGCGC 3538

QY 223 CATCTGCCAGAGAGCTGTGACCACTTC 253
|||||
3539 CACCTCCAGAGAGACTGGCGCGCGCTCC 3569

Db

RESULT 14
G67139 448 bp DNA linear STS 18-SEP-2000
LOCUS IGFBP1 X1.1 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION
ACCESSION G67139
VERSION G67139.1 GI:10186730
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 448)
Cox,D.G., Bollnot,C. and Canzian,F.
Genome Survey
Unpublished (2000)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Federico Canzian
Genome Analysis Group
International Agency for Research on Cancer
150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France
Tel: +33-4-72738698
Fax: +33-4-72738388
Email: canzian@arc.fr
Primer A: TGCACCTACGCAAAACAACCTT
Primer B: GGGTGACCTCCAGACAG
STS size: 448
Protocol:
Template: 50 ng
Primer: 4 um each
dNTPs: 2 mM each
MgCl2: 1.5-2.5 mM
Tag: 0.05 units
Total Vol: 25 ul
Buffer: 1.5-2.5 mM
MgCl2: 50 mM
KCl: 20 mM
Tris-HCl: 8.4.
pH:

FEATURES
source 1. 448
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind 1. 448
primer_bind 1. 20
complement(431. 448)
ORIGIN

Query Match 22.5%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 4.5e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTACACGCTGCTGCAATCAT 112
DB 3 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGCGCGCTGCCAATCAT 62
QY 113 TAACCC-----GTGCTGCCGAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 162
DB 63 TAACCTCTGTGTCAGAGTGCGCGGCTGTGCTTTATTAAGTGTGCGCTGTGTCCAGC 122
QY 163 CAGCATGTCTCACTGCGCGCGAGACACAAACCGAGCATTTGAACACTGC-ACACGG 221
DB 123 GAGCATCGGCGACCGCATCC-----CATCCAGCGAGCATCTGCCCGCGCGCGCG 174
QY 222 CCATCTGCCGAGAGAGCTGTGACCAACACTTC 253
DB 175 CCACCTCCGAGAGAGCACTGGCGACCGCTCC 206

RESULT 15
AX409747 6128 bp DNA linear PAT 14-JUN-2002
LOCUS AX409747
DEFINITION Sequence 2194 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:21442452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
Alvarez, C., Horne, D., Perez-da-Silva, S. and Vockley, J.G.
AUTHORS
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-ABR-2002;
GENE LOGIC INC (US)
FEATURES
source 1. 6128
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN

Query Match 22.5%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 4.5e-07;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGATCTTACACGCTGCTGCAATCAT 112
DB 473 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGCGCGCTGCCAATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCCCTTCATTAAGCCCTGGGTATGGCCAGC 162
DB 533 TAACCTCTGTGTCAGAGTGCGCGGCTGTGCTTTATTAAGTGTGCGCTGTGTCCAGC 592
QY 163 CAGCATGTCTCACTGCGCGCGAGACACAAACCGAGCATTTGAACACTGC-ACACGG 221
DB 593 GAGCATCGGCGACCGCATCC-----CATCCAGCGAGCATCTGCCCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGAGCTGTGACCAACACTTC 253

DB 645 CCACCTCCGAGAGAGCACTGGCGACCGCTCC 676

Search completed: September 1, 2005, 03:26:49
Job time : 1685.14 secs

This Page Blank (uspto)

[illegible]

25	67.4	25.0	488	1	AA674302	AA674302	VP6G910. r
26	65.4	24.2	801	7	CK472246	CK472246	AGENCOURT
27	64.4	23.9	840	7	CK473709	CK473709	AGENCOURT
28	62.8	23.3	759	1	AA105355	AA105355	mp37d09. r
29	61.8	22.9	259	2	BB604790	BB604790	BB604790
30	61.6	22.8	587	2	AW916227	AW916227	EST347531
31	60.8	22.5	1589	3	CR621807	CR621807	full-1eng
32	60.8	22.5	1601	3	CR650377	CR650377	full-1eng
33	59	21.9	5451	3	AA060360	AA060360	mj67a12. r
34	43	15.9	659	8	AZ840793	AZ840793	2M0138D02
35	38.8	14.4	522	2	AA655007	AA655007	105551. MA
36	37.2	13.8	522	2	AA655007	AA655007	105551. MA
37	36.8	13.6	428	2	BF829782	BF829782	MR3-HN006
38	36.8	13.6	849	9	CNS02065	CNS02065	Tetradon
39	36.6	13.6	632	2	CNS007MH	CNS007MH	Drocephal1
40	35	13.0	480	5	BX282084	BX282084	787256
41	35	13.0	480	5	BX282084	BX282084	787256
42	35	13.0	509	1	AA443181	AA443181	z98h10. r
43	35	13.0	532	4	BG283072	BG283072	602406595
44	35	13.0	533	4	BG388790	BG388790	602414492
45	35	13.0	561	4	BM772026	BM772026	K-BEST0056

ALIGNMENTS

RESULT	1
LOCUS	BF236974
DEFINITION	BF236974 1013 bp mRNA linear EST 14-NOV-2000 602026605f1 NCI CGAP_Li9 Mus musculus cDNA clone IMAGE:416188 5'
ACCESSION	BF236974
VERSION	BF236974
KEYWORDS	EST.
SOURCE ORGANISM	Mus musculus (house mouse)
REFERENCE	NH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Straubeberg, Ph.D.
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE	(bases 1 to 1013)
JOURNAL COMMENT	

FEATURES

```

FEATURES
    source
        cdna_library Arrayed by: The I.M.A.G.E. Consortium (LNL1)
        dna_sequencing by: Incyte Genomics, Inc.
        clone_distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL1 at:
            http://image.lnl.gov
        plate: LLM9443 row: j column: 02
        high_quality sequence stop: 581.
        location/Qualifiers
            1..1013
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGR:4161889"
            /lab_host="DH10B (rti phage-resistant)"
            /clone_id="NC1_CGAP_1.9"
            /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: Nci1;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.9 kb. Constructed by Life
                Technologies. Note: this is a Nci_CGAP library."

```

ORIGIN

Query Match	36.4%	Score 98.2	DB 2	Length 1013
Best Local Similarity	92.8%	Pred. No. 1.5e-18		
Matches 103	Conservative	0	Mismatches 8	Indels 0
			Gaps 0	
155	TGGCCGACGATGTTCACCTCCCGCCGACACAAACCGACGACATTGAACCTG	214		

Db 9 TGGGAGGCGAGATGGTCCACTGCGCGGAGACACACCCAGGAGATTGAACACTG 68

QY 215 CACAGGCGCATTCGCCAGAGAGCTGTACCAACCACTTCGCTACTACTGTA 265

Db 69 CACAGGCGCGTGTGCCAGAGAGCTGTGACCACTTGCCTACTACTACTA 119

RESULT 2
BY783538

LOCUS 310 bp mRNA linear EST 23-MAR-2004

DEFINITION BY783538 RIKEN full-length enriched, 17.5 days embryo whole body

ACCESSION BY783538

VERSION BY783538.1 GI:39710177

KEYWORDS EST

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus. 1 (bases 1 to 310)

AUTHORS Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavani, W., Aldinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakuchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

TITLE Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia

JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)

MEDLINE 22703353

PUBMED 12819125

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for further details.

FEATURES
source location/Qualifiers
1..310
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U930176D05"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_idb="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN

Query Match 32.5%; Score 87.8; DB 6; Length 310;
Best Local Similarity 92.9%; Pred. No. 1.6e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 167 ATGTGCACTGCGCGGAGACACAACCCAGGAGATTGAACATGCAACGGCCATC 226

Db 2 ATGTGCACTGCGCGGAGACACAACCCAGGAGATTGAACATGCAACGGCCGTC 61

QY 227 TGCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTGTA 265

Db 62 TGCCAGAGAGCTGTGACCAACCACTTGCCTACTACTATCTA 100

RESULT 3
BY794229

LOCUS 331 bp mRNA linear EST 23-MAR-2004

DEFINITION BY794229 RIKEN full-length enriched, 17.5 days embryo whole body

ACCESSION BY794229

VERSION BY794229.1 GI:39720868

KEYWORDS EST

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus. 1 (bases 1 to 331)

AUTHORS Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavani, W., Aldinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakuchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

TITLE Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia

JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)

MEDLINE 22703353

PUBMED 12819125

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for further details.

FEATURES
source location/Qualifiers
1..331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="U930292H24"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_idb="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN

Query Match 32.5%; Score 87.8; DB 6; Length 331;
Best Local Similarity 92.9%; Pred. No. 1.6e-15;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 167 ATGTGCACTGCGCGGAGACACAACCCAGGAGATTGAACATGCAACGGCCATC 226

Db 2 ATGTGCACTGCGCGGAGACACAACCCAGGAGATTGAACATGCAACGGCCGTC 61

QY 227 TGCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTGTA 265

Db 62 TGCCAGAGAGCTGTGACCAACCACTTGCCTACTACTATCTA 100

RESULT 4
BY71317

LOCUS	BY7171317	354 bp	mRNA	linear	EST 23-MAR-2004
DEFINITION	BY7171317 RIKEN full-length enriched, 17.5 days embryo whole body				
ACCESSION	Mus musculus cDNA clone U930052G15 5', mRNA sequence.				
VERSION	BY7171317				
KEYWORDS	BY7171317.1 GI:39697955				
SOURCE	EST.				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	Mus musculus				
AUTHORS	Emura/Yota; Metazawa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 354)				
TITLE	Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Saeki, D., Bono, H., Kondo, S., Sugihara, Y., Salto, R., Oseato, N., Fukuda, S., Sato, K., Watabiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Guetlich, S., Beisel, K., Pavan, W., Aldini, V., Nakagawa, A., Heid, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Henschi, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J., and Hayashizaki, Y.				
JOURNAL	Targeting a complex transcripome: the construction of the mouse full-length cDNA encyclopedia				
MEDLINE	Genome Res. 13 (6B), 1273-1289 (2003)				
PUBMED	22703353				
COMMENT	12819125				
	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216				
	Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.				
FEATURES	location/Qualifiers				
source	1..354				
	/organism="Mus musculus"				
	/mol_type="mRNA"				
	/strain="C57BL/6J"				
	/db_xref="taxon:10090"				
	/clone="U930052G15"				
	/tissue_type="whole body"				
	/dev_stage="17.5 days embryo"				
	/clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"				
ORIGIN					
	Query Match 32.5%; Score 87.8; DB 6; Length 354;				
	Best Local Similarity 92.9%; Pred. No. 1.7e-15;				
	Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;				
QY	167 ATGTGTCACCTGCCCGCAGACAAACCCGACGACATTGGAACATGCACACGGCCATC 226				
DB	2 ATGTGTCACCTGCCCGCAGACAAACCCGACGACATTGGAACATGCACACGGCCATC 61				
QY	227 TGCCAGAGACTGTGACCACTTCGCGTACTAGCTA 265				
DB	62 TGCCAGAGACTGTGACCACTTCGCGTACTAGCTA 100				
RESULT 5	AT785818	401 bp	mRNA	linear	EST 02-JUL-1999
LOCUS	u178h05.v1 Sugano mouse liver b1a Mus musculus cDNA clone				
DEFINITION	IMAGE:188565 5', similar to gb:X61579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.				
ACCESSION	AT785818				

VERSION	A1785618.1	GI:5333534
KEYWORDS	EST,	Mus musculus (house mouse)
SOURCE		Eukaryote
ORGANISM		Mus musculus
REFERENCE		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
AUTHORS		1 (bases 1 to 401) Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepeco,M., Theising,B., Allen,M., Bowers,Y., Persson,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Riter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE		The Mashu-NCI Mouse EST Project 1999
JOURNAL		Unpublished (1999)
COMMENT		Other ESTs: u178h05.x1 Contact: Marra M/Mashu-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@atcson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. MG1:972893
FEATURES		Seq primer: custom primer used High quality sequence stop: 126. Location/Qualifiers
SOURCE		1..401 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL." /db_xref="taxon:10090" /clone="IMAGE:1888569" /sex="female" /dev_stage="adult" /lab_host="DH10B" /clone_1lb="Sugano mouse liver mlA" /note="Organ: liver; Vector: pME18S-FLJ; Site: 1: DraIII (CACTGTG); Site 2: DraII (CACCATGG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTTCCTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TGTTGGCTACTG), digested and cloned into distinct DraIII sites of the pME18S-FLJ vector (5' site CACTGTGG, 3' site CACCATGG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTTAAGAAGTCGC and 3' end primer CGACTGTGACGCTCGACACA."
ORIGIN		
Query Match	32.1%;	Score 86.0; DB 1; Length 401;
Best Local Similarity	92.9%;	Pred. No. 3,4e-15;
Matches	91;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY	168	TGTCGCACTGCCGCGGAGACACAAACCGAGCGATTGAACATGACACGGCCATC 227
DB	1	TGGTCACCTGCCCGCGGAGACACACCACCGAGCATTAACACATGACACGCGCCTC 60
OY	228	GCCCAGAGAGCTGTGACACCACTTCCGCTACTAGCTA 265
DB	61	GCCAGAGAGCTGTGACACCACTTCCGCTACTAGCTA 98
RESULT 6		
LOCUS	A1196314	480 bp mRNA linear EST 14-OCT-1998
DEFINITION		u171a07.v1 Sugano mouse liver mlA Mus musculus cDNA clone IMAGE:1887828 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HDMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellendberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Watson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973152
 FEATURES
 source
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCGAGCTCGACACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 3.6e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGGTCCACTGCGCCCGGAGACACAAACCAGAGCATTTGAACACTGCACAGCGCATCT 227
 DB 1 TGGTCCACTGCGCCCGGAGACACACACCCAGAGCATTTGAACACTGCACAGCGCATCT 60
 QY 228 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION u189f09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:W5316 rat insulin-like GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146 GI:4444281
 VERSION A1530146.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 FEATURES
 source
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCGAGCTCGACACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 3.9e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGGTCCACTGCGCCCGGAGACACAAACCAGAGCATTTGAACACTGCACAGCGCATCT 227
 DB 1 TGGTCCACTGCGCCCGGAGACACACACCAGAGCATTTGAACACTGCACAGCGCATCT 60
 QY 228 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION u311e07.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579 M.musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 765)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930344
 Seq primer: custom primer used
 High quality sequence stop: 395.
 FEATURES
 source
 1..765
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1481988"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI08"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FU3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 765;
 Best Local Similarity 92.9%; Pred. No. 4e-15; 7; Indels 0; Gaps 0;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGTTCACCTGCCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCACTT 227
 DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCACTT 60
 QY 228 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 9
 A1530313 795 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530313
 DEFINITION u191f01.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
 VERSION A1530313.1 GI:4444448
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 785)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974109
 Seq primer: custom primer used
 High quality sequence stop: 459.
 FEATURES
 source
 1..785
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889785"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DHI08"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18S-FU3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCAGCTCGAGCACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 785;
 Best Local Similarity 92.9%; Pred. No. 4e-15; 7; Indels 0; Gaps 0;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 168 TGTTCACCTGCCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACACGGCCACTT 227
 DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGAGCATTTGAACACTGCACACGGCCACTT 60
 QY 228 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 265
 DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 10
 A1529939 799 bp mRNA linear EST 18-MAR-1999
 LOCUS A1529939
 DEFINITION u167c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 IMAGE:1889392 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1529939 GI:4444074
 VERSION A1529939.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 799)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGB Consortium (info@image.lnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 FEATURES
 source location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pMB18-FL3; Site 1: Draili (CAGCTGTG); Site 2: Draili (CAGCAGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a Draili adaptor [TGTGCGCTACTG]; digested and cloned into distinct Draili sites of the pMB18-FL3 vector (5' site CAGCTGTG, 3' site CAGCAGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTMAAAGCTGCG and 3' end primer CGACTGCGAGCTCGAGCACA."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 4e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 168 TGTGTCACATGCGCGCGAGACACAAACCCAGAGCATTTGAACATCGACACAGGCCATCT 227
 |||||||
 Db 1 TGTGTCACATGCGCGCGAGACACACCCAGAGCATTTGAACATCGACACAGGCCGCTCT 60
 |||||||
 Oy 228 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 265
 |||||||
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTATCTA 98
 |||||||
 RESULT 11 811 bp mRNA linear EST 09-MAR-2004
 CA478518
 LOCUS CA478518
 DEFINITION AGENCOURT.10789306 NIH_MGC_152 Mus musculus cDNA clone
 IMAGE:6766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LMK00138 row: m column: 23
 High quality sequence stop: 536.
 FEATURES
 source location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: atcP2; Site 2: atcP1; cDNA made by oligo-dT with atcB2 site and directionally cloned. Priming sequence:
 5'-TTTCCTGCGAGCGCGCGACCACTTGTGACAGAGCGGCTTTTCTTTTCTTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen, library amplification 16 cycles. Library constructed by Mark Bittinger in the Bradfield Laboratory (McArdle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 32.1%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 4e-15;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 168 TGTGTCACATGCGCGCGAGACACAAACCCAGAGCATTTGAACATCGACACAGGCCATCT 227
 |||||||
 Db 39 TGTGTCACATGCGCGCGAGACACACCCAGAGCATTTGAACATCGACACAGGCCGCTCT 98
 |||||||
 Oy 228 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 265
 |||||||
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTCTA 136
 |||||||
 RESULT 12 848 bp mRNA linear EST 02-JUL-1999
 A1790802
 LOCUS A1790802
 DEFINITION uk28b10.y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 IMAGE:1970299 5', similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802 GI:5338518
 VERSION A1790802.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 848)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28b10.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039

FEATURES

High quality sequence stop: 514.
Location/Qualifiers
1..848

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse kidney m1a"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGG); Site 2: DraIII (CACTGTGG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGG, 3' site CACTGTGG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 32.1%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 4e-15;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCGCGCGAGACACAAACCAGGAGCATTTGAACACTGCACAGCGCATCT 227

DB 1 TGGTCCACTGCGCGCGAGACACACACCAGGAGCATTTGAACACTGCACAGCGCATCT 60

QY 228 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 265

DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 13
A1528304 852 bp mRNA linear EST 18-MAR-1999
A1528304

LOCUS u195g10.y1 Sugano mouse liver m1a Mus musculus cDNA clone
DEFINITION IMAGE:1890210.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH

FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1528304
VERSION A1528304.1 GI:4442439

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 852)

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534

FEATURES

High quality sequence stop: 478.
Location/Qualifiers
1..852

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver m1a"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGG); Site 2: DraIII (CACTGTGG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGG, 3' site CACTGTGG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGCTCTAAAGCTGG and 3' end
primer CGACTCGAGCTCGAGCACA."

ORIGIN

Query Match 32.1%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 4e-15;

Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCGCGCGAGACACAAACCAGGAGCATTTGAACACTGCACAGCGCATCT 227

DB 1 TGGTCCACTGCGCGCGAGACACACACCAGGAGCATTTGAACACTGCACAGCGCATCT 60

QY 228 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 265

DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998
A1196154

LOCUS u195d08.y1 Sugano mouse liver m1a Mus musculus cDNA clone
DEFINITION IMAGE:1887665.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH

FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1196154
VERSION A1196154.1 GI:3748760

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 605)

Marra, M., Hillier, L., Kucaba, T., Allen, M., Bowers, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMNI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMII Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:971987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

```
1..605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TTTGGCTTACTGG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer GCACCTGCAGCTCGACACA."
```

ORIGIN

Query Match

Best Local Similarity 31.6%; Score 85.2; DB 1; Length 605;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 168 TGGTCCACTGCCCGCGGAGACACAAACCAGAGCATTTGAACACTGCACAGCGCATCT 227
|||||
DB 1 TGGTCCACTGCCCGCGGAGACACACACCCAGAGATTGAACACTGCACAGCGTCTCT 60
|||||
QY 228 GCCCAGAGAGCTGTGAACCACTTCGCTACTAGCTA 265
|||||
DB 61 GCCCAGAGAGCTGTGAACCACTTCGCTACTATCTA 98
|||||

RESULT 15

AT785039

LOCUS

DEFINITION

u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone.
IMAGE:1888018 5' similar to gp:X81579 M.musculus mRNA for
innulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AT785039 380 bp mRNA linear EST 02-JUL-1999
u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone.
IMAGE:1888018 5' similar to gp:X81579 M.musculus mRNA for
innulin-like growth factor binding (MOUSE);, mRNA sequence.

REFERENCE

AUTHORS

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theisling, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

```
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor (TTTGGCTTACTGG), digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTGTGCTCTAAAGCTGCG and 3' end
primer GCACCTGCAGCTCGACACA."
```

ORIGIN

Query Match

Best Local Similarity 30.4%; Score 82.2; DB 1; Length 380;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 171 TCCACTGCCCGCGGAGACACAAACCAGAGCATTTGAACACTGCACAGCGCATCTGCGC 230
|||||
DB 1 TCCACTGCCCGCGGAGACACACACCCAGAGCATTTGAACACTGCACAGCGCGCTCGCC 60
|||||
QY 231 CAGAGAGCTGTGAACCACTTCGCTACTAGCTA 265
|||||
DB 61 CAGAGAGCTGTGAACCACTTCGCTACTATCTA 95
|||||

Search completed: September 1, 2005, 06:56:48
Job time : 2054.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 786.359 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-3

Perfect score: 270

Sequence: 1 catggcgccagcgaggcgactc.....ttccgctactagctacggcgc 270

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	270	100.0	270	US-09-972-916A-3	Sequence 3, Appli
2	270	100.0	423	US-09-972-916A-6	Sequence 6, Appli
3	219.4	81.3	321	US-09-972-916A-4	Sequence 4, Appli
4	219.4	81.3	372	US-09-972-916A-5	Sequence 5, Appli
5	219	81.1	219	US-09-972-916A-2	Sequence 2, Appli
6	76.4	28.3	1500	US-09-917-800A-1608	Sequence 1608, Ap
7	76.4	28.3	1500	US-10-388-934-5	Sequence 5, Appli

8	76.4	28.3	1500	17	US-10-191-803-73	Sequence 73, Appli
9	76.4	28.3	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
10	60.8	22.5	6128	9	US-09-880-107-2293	Sequence 2393, Ap
11	60.8	22.5	6128	22	US-10-756-149-1484	Sequence 1484, Ap
12	60.8	22.5	9173	22	US-10-893-315-126	Sequence 126, App
13	60.8	22.5	9174	22	US-10-893-315-160	Sequence 160, App
14	51	18.9	51	9	US-09-972-916A-1	Sequence 1, Appli
15	51	18.9	321	9	US-09-972-916A-4	Sequence 4, Appli
16	51	18.9	372	9	US-09-972-916A-5	Sequence 5, Appli
17	51	18.9	13011	17	US-10-388-934-36	Sequence 36, Appli
18	51	18.9	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
19	37.8	14.0	7061	15	US-10-311-655-970	Sequence 970, App
20	37.8	14.0	7061	17	US-10-221-613-148	Sequence 148, App
21	36.6	13.6	31703	17	US-10-085-117-172	Sequence 172, App
22	36.4	13.5	316778	13	US-10-027-632-174961	Sequence 174961, Sequence 174961,
23	36.4	13.5	316778	17	US-10-027-632-174961	Sequence 174961, Sequence 174961,
24	35.2	13.0	1687	20	US-10-739-930-5461	Sequence 5461, Ap
25	35	13.0	769	13	US-10-027-632-164336	Sequence 164336, Sequence 164336,
26	35	13.0	769	17	US-10-027-632-164336	Sequence 164336, Sequence 164336,
27	35	13.0	3895	14	US-10-011-585A-76	Sequence 76, Appli
28	34	12.6	49979	19	US-10-741-601-5746	Sequence 5746, Appli
29	34	12.6	49979	21	US-10-741-600-17905	Sequence 17905, A
30	33.4	12.4	1806	19	US-10-250-615-33	Sequence 33, Appli
31	33.4	12.4	2667	17	US-10-108-260A-2239	Sequence 2239, Ap
32	32.4	12.0	1389	19	US-10-437-963-79162	Sequence 79162, A
33	32.4	12.0	1584	19	US-10-437-963-79185	Sequence 79185, A
34	32	11.9	76180	19	US-10-332-281-492	Sequence 492, App
35	31.6	11.7	1129	20	US-10-343-903-31	Sequence 31, Appli
36	31.6	11.7	1288	15	US-10-017-161-2017	Sequence 2017, Ap
37	31.6	11.7	1288	17	US-10-292-798-1663	Sequence 1663, Ap
38	31.6	11.7	1348	17	US-10-120-988-404	Sequence 404, App
39	31.6	11.7	1513	18	US-10-072-012-29	Sequence 29, Appli
40	31.6	11.7	1555	10	US-09-764-872-581	Sequence 581, App
41	31.6	11.7	11750	19	US-10-367-094-169	Sequence 169, App
42	31.4	11.6	360	20	US-10-425-115-28332	Sequence 28332, A
43	31.4	11.6	1134	13	US-10-282-122A-33251	Sequence 33251, A
44	31.4	11.6	16554	17	US-10-087-192-391	Sequence 391, App
45	31.2	11.6	668	20	US-10-363-345A-40493	Sequence 40493, A

ALIGNMENTS

RESULT 1
US-09-972-916A-3
Sequence 3, Application US/09972916A
Patent No. US20020107198A1
GENERAL INFORMATION:
APPLICANT: Thule, Peter M.
TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
FILE REFERENCE: US 1292/01 (VA)
CURRENT APPLICATION NUMBER: US/09/972, 916A
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/239, 113
PRIOR FILING DATE: 2000-10-11
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 3
LENGTH: 270
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 100.0%; Score 270; DB 9; Length 270;
Best Local Similarity 100.0%; Pred. No. 3.2e-86;
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTGTGAGCCCAAGTATCAACAGCA 60
DB 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTGTGAGCCCAAGTATCAACAGCA 60
QY 61 AAACAATTATTTTGAACACAGCGGATCTCTAGACACGCTGCCCTGACATCATTAACCGGT 120

```

Db      61 AAAAACAATTATTTGAACACGGGATCTTAGACGCTGCCCTGACAATCATTAACCGT 120
        |||
Qy      121 GCTGCGAGACCAACCCCTTATAGAGCCCTGGGTATATGCGACAGATGCTCACTGCC 180
        |||
Db      121 GCTGCGAGACCAACCCCTTATAGAGCCCTGGGTATATGCGACAGATGCTCACTGCC 180
        |||
Qy      181 GCCGAGACCAAAACCCAGGAGCATTAACACTGACAGGCGCATCTGCCAGAGAGCTG 240
        |||
Db      181 GCCGAGACCAAAACCCAGGAGCATTAACACTGACAGGCGCATCTGCCAGAGAGCTG 240
        |||
Qy      241 TGACCAACCACTTCGCTACTAGTACCGCG 270
        |||
Db      241 TGACCAACCACTTCGCTACTAGTACCGCG 270
        |||

```

RESULT 2

```

US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6

```

Query Match 100.0%; Score 270; DB 9; Length 423;
 Best Local Similarity 100.0%; Pred. No. 3.7e-86;
 Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 CATGGCGCAGCGGGGACTCCCGTGGTCTCTGACTGTGCGCCCAAGTGTATCACAAGA 60
        |||
Db      154 CATGGCGCAGCGGGGACTCCCGTGGTCTCTGACTGTGCGCCCAAGTGTATCACAAGA 213
        |||
Qy      61 AAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAATCATTAACCGT 120
        |||
Db      214 AAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAATCATTAACCGT 273
        |||
Qy      121 GCTGCGAGACCAACCCCTTATAGAGCCCTGGGTATATGCGACAGATGCTCACTGCC 180
        |||
Db      274 GCTGCGAGACCAACCCCTTATAGAGCCCTGGGTATATGCGACAGATGCTCACTGCC 333
        |||
Qy      181 GCCGAGACCAAAACCCAGGAGCATTAACACTGACAGGCGCATCTGCCAGAGAGCTG 240
        |||
Db      334 GCCGAGACCAAAACCCAGGAGCATTAACACTGACAGGCGCATCTGCCAGAGAGCTG 393
        |||
Qy      241 TGACCAACCACTTCGCTACTAGTACCGCG 270
        |||
Db      394 TGACCAACCACTTCGCTACTAGTACCGCG 423
        |||

```

RESULT 3

```

US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113

```

```

; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4

```

Query Match 81.3%; Score 219.4; DB 9; Length 321;
 Best Local Similarity 99.5%; Pred. No. 4.7e-68;
 Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      50 TATCAAGCAAAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAT 109
        |||
Db      101 TGTCAAGCAAAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAT 160
        |||
Qy      110 CATTAAACCGGTGCTGCCGAGCAGCCCTTCATTAAGGCCCTGGTATGCGCAGCAGATG 169
        |||
Db      161 CATTAAACCGGTGCTGCCGAGCAGCCCTTCATTAAGGCCCTGGTATGCGCAGCAGATG 220
        |||
Qy      170 GTCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 229
        |||
Db      221 GTCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 280
        |||
Qy      230 CCAGAGACTGTGACCAACCACTTCGCTACTAGTACCGCG 270
        |||
Db      281 CCAGAGACTGTGACCAACCACTTCGCTACTAGTACCGCG 321
        |||

```

RESULT 4

```

US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

```

Query Match 81.3%; Score 219.4; DB 9; Length 372;
 Best Local Similarity 99.5%; Pred. No. 5e-68;
 Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      50 TATCAAGCAAAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAT 109
        |||
Db      152 TGTCAAGCAAAACAACTTATTTGAACACGGGATCTTAGACGCTGCCCTGACAT 211
        |||
Qy      110 CATTAAACCGGTGCTGCCGAGCAGCCCTTCATTAAGGCCCTGGTATGCGCAGCAGATG 169
        |||
Db      212 CATTAAACCGGTGCTGCCGAGCAGCCCTTCATTAAGGCCCTGGTATGCGCAGCAGATG 271
        |||
Qy      170 GTCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 229
        |||
Db      272 GTCACTGCCCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACGCGCATCTGC 331
        |||
Qy      230 CCAGAGACTGTGACCAACCACTTCGCTACTAGTACCGCG 270
        |||
Db      332 CCAGAGACTGTGACCAACCACTTCGCTACTAGTACCGCG 372
        |||

```



```
RESULT 5
US-09-972-916A-2
; Sequence 2, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
US-09-972-916A-2

Query Match      81.1%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 5.8e-68;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      52 TCACAAGCAAAACCTATTATTTTGAACACGCGGAGATCTTAGACGCGTCCCTGACATCA 111
DB      1 TCACAGCAAAACCAACTTATTATTGACACGCGGATCTTAGACGCTGCTGACATCA 60
QY      112 TTAACCGGTGCTGCGGAGCGAGCCCTTCTATTAAGCCCTGGATAGCCAGCAGCATGCT 171
DB      61 TTAACCGGTGCTGCGGAGCGAGCCCTTCTATTAAGCCCTGGATAGCCAGCAGCATGCT 120
QY      172 CCACGTCCCGCCGAGACACAAACCCAGCAGCATTTGAACACTGACACAGCCCATCTGCC 231
DB      121 CCACGTCCCGCCGAGACACAAACCCAGCAGCATTTGAACACTGACACAGCCCATCTGCC 180
QY      232 AGAGAGCTGTGACACCACTTCCGCTACTAGCTAGCCGC 270
DB      181 AGAGAGCTGTGACACCACTTCCGCTACTAGCTAGCCGC 219

RESULT 6
US-09-917-800A-1608
; Sequence 1608, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
```

```
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1608
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match      28.3%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCAGCATTTGAACACTGACACAGGCGCATCTGCCAGAGAGCTGTGACCA 247
DB      8 CACAAACCCAGCAGCATTTGAACACTGACACAGGCGCATCTGCCAGAGAGCTGTGACCA 67
QY      248 CACTTCCGCTACTAGCTA 265
DB      68 CACTTCCGCTACTAGCTA 85

RESULT 7
US-10-388-934-5
; Sequence 5, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boese, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547A1 way rat)
US-10-388-934-5

Query Match      28.3%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      188 CACAAACCCAGCAGCATTTGAACACTGACACAGGCGCATCTGCCAGAGAGCTGTGACCA 247
DB      8 CACAAACCCAGCAGCATTTGAACACTGACACAGGCGCATCTGCCAGAGAGCTGTGACCA 67
QY      248 CACTTCCGCTACTAGCTA 265
DB      68 CACTTCCGCTACTAGCTA 85

RESULT 8
US-10-191-803-73
; Sequence 73, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
```

FILE REFERENCE: 44921-5090US
CURRENT APPLICATION NUMBER: US/10/191,803
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: US 60/303,819
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/305,623
PRIOR FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: US 60/369,351
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/377,611
PRIOR FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 1140
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 73
LENGTH: 1500
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_013144
US-10-191-803-73

Query Match 28.3%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
DB 8 CACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67

QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 9
US-10-152-319A-1613
Sequence 1613, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgins, Brandon
APPLICANT: Casle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1613

LENGTH: 1500
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match 28.3%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 188 CACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 247
DB 8 CACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 67

QY 248 CACTTCCGCTACTAGCTA 265
DB 68 CACTTCCGCTACTAGCTA 85

RESULT 10
US-09-880-107-2393
Sequence 2393, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Scherf, Joseph G.
APPLICANT: Scherf, Iwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2393
LENGTH: 6128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
US-09-880-107-2393

Query Match 22.5%; Score 60.8; DB 9; Length 6128;
Best Local Similarity 66.5%; Pred. No. 8.9e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCCAC 112
DB 473 CACTTCCGCTACTAGCTA 532

QY 113 TAACCC-----GTGCTGCCAGCAGCCCTTATAGAGCGCTGATGAGCCAC 162
DB 533 TAACCTCTGTGTAAGTGGCGGCGCTGTGCTTTAAGTGGCGCTGTGCAAGC 592

QY 163 CAGCATGTCCTACTGCGCCCGGAGACAAACCCAGGAGCATTTGAACACTGC-ACAAG 221
DB 593 GAGCATGGCGCACCGCATCC-----CATCAGAGAGCATGTGCGCGCGCGCG 644

QY 222 CCATCTGCCAGAGAGCTGTGACCCACCACTTC 253
DB 645 CCACCTTCCAGAGAGCATGTGCGCACCGCTCC 676

RESULT 11
US-10-756-149-1484
Sequence 1484, Application US/10756149
Publication No. US20050181375A1
GENERAL INFORMATION:

```

; APPLICANT: Aziz, Nareasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: Patent version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-756-149-1484

Query Match      22.5%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 8-9e-11;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTACACGCTGCTGACATCAT 112
DB 473 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCAATCAT 532
QY 113 TAACCC-----GTGCTGCCGAGCCGCTTCAATTAAGCCCTGGGTATGGCCAGC 162
DB 533 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTCTTATTAAGTGGCGCTGTCTCCAGC 592
QY 163 CAGCATGTGTCACTGCCCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 593 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 644
QY 222 CCATCTGCCGAGAGCTGTGACCACTTC 253
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 12
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-126

Query Match      22.5%; Score 60.8; DB 22; Length 9173;
Best Local Similarity 66.5%; Pred. No. 1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTACACGCTGCTGACATCAT 112
DB 1877 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCAATCAT 1936
QY 113 TAACCC-----GTGCTGCCGAGCCGCTTCAATTAAGCCCTGGGTATGGCCAGC 162
DB 1937 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTCTTATTAAGTGGCGCTGTCTCCAGC 1996
QY 163 CAGCATGTGTCACTGCCCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 1997 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 2048

```

```

QY 222 CCATCTGCCGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGCACTGGCCACCGCTCC 2080

RESULT 13
US-10-893-315-160
; Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-160

Query Match      22.5%; Score 60.8; DB 22; Length 9174;
Best Local Similarity 66.5%; Pred. No. 1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAAGCAAAACAACTTATTTTGAACACGGGATCTTACACGCTGCTGACATCAT 112
DB 1877 CACTGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGTGGCGGCTGCCAATCAT 1936
QY 113 TAACCC-----GTGCTGCCGAGCCGCTTCAATTAAGCCCTGGGTATGGCCAGC 162
DB 1937 TAACCTCTGCTGCAAGTGGCGGCGCTGTGCTCTTATTAAGTGGCGCTGTCTCCAGC 1996
QY 163 CAGCATGTGTCACTGCCCGCCGAGACACAAACCGACGAGATTGAACACTGC-ACACGG 221
DB 1997 GAGCATGCGGCCACCGCCATCC-----CATCCAGCGAGCATCTGCGCGCGCGCGCG 2048
QY 222 CCATCTGCCGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGCACTGGCCACCGCTCC 2080

RESULT 14
US-09-972-916A-1
; Sequence 1, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 1
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-972-916A-1

Query Match      18.9%; Score 51; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCGCCAGTGTGA 51
 |||||
 Db 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCGCCAGTGTGA 51

RESULT 15
 US-09-972-916A-4/c
 ; Sequence 4, Application US/09972916A
 ; Patent No. US20020107198A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thule, Peter M.
 ; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
 ; FILE REFERENCE: US 1292/01 (VA)
 ; CURRENT APPLICATION NUMBER: US/09/972,916A
 ; CURRENT FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/239,113
 ; PRIOR FILING DATE: 2000-10-11
 ; NUMBER OF SEQ ID NOS: 6
 ; SEQ ID NO 4
 ; LENGTH: 321
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthesized
 US-09-972-916A-4

Query Match 18.9%; Score 51; DB 9; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCGCCAGTGTGA 51
 |||||
 Db 102 CATGGCGCAGCGGCACTCCCGTGGTTCCTGGAAGCTGAGCCGCCAGTGTGA 52

Search completed: September 1, 2005, 16:44:18
 Job time : 793.359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 85.2717 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916b-3

Perfect score: 270
Sequence: 1 catggggcagcagggcagc.....ttcgcctactagctagccgc 270

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.8	22.5	9173	4 US-09-949-001-30	Sequence 30, Appl
2	60.8	22.5	9174	4 US-09-949-001-36	Sequence 36, Appl
3	51	18.9	194	3 US-08-945-140-1	Sequence 1, Appl
4	51	18.9	13011	2 US-08-791-849A-14	Sequence 14, Appl
5	36.6	13.6	15108	4 US-09-949-016-11786	Sequence 11786, A
6	36.6	13.6	15108	4 US-09-949-016-117205	Sequence 117205, A
7	35	13.0	601	4 US-09-949-016-13194	Sequence 13194, A
8	35	13.0	601	4 US-09-949-016-13194	Sequence 13194, A
9	35	13.0	2393	4 US-09-949-016-2153	Sequence 2153, Ap
10	35	13.0	2408	4 US-09-949-016-74	Sequence 74, Appl
11	35	13.0	7561	4 US-09-949-016-11816	Sequence 11816, A
12	35	13.0	7562	4 US-09-949-016-13895	Sequence 13895, A
13	32.2	11.9	4447	2 US-08-304-309-3	Sequence 3, Appl
14	32.2	11.9	4447	3 US-08-991-942-3	Sequence 3, Appl
15	32	11.9	36180	4 US-09-949-016-11745	Sequence 11745, A
16	32	11.9	36181	4 US-09-949-016-16163	Sequence 16163, A
17	31.6	11.7	1348	4 US-09-774-528-404	Sequence 404, App
18	31.6	11.7	99748	4 US-09-949-016-11990	Sequence 11990, A
19	31.6	11.7	99749	4 US-09-949-016-16518	Sequence 16518, A
20	31.2	11.6	364	4 US-09-270-767-23006	Sequence 23006, A
21	31.2	11.6	601	4 US-09-949-016-159692	Sequence 159692, A
22	31.2	11.6	601	4 US-09-949-016-159693	Sequence 159693, A
23	31.2	11.6	761	4 US-09-270-767-13109	Sequence 13109, A
24	31	11.5	1888	4 US-09-949-016-1341	Sequence 1341, Ap
25	31	11.5	13145	4 US-09-949-016-13083	Sequence 13083, Ap
26	30.8	11.4	1800	4 US-09-252-991A-3891	Sequence 3891, Ap
27	30.8	11.4	1824	4 US-09-252-991A-4025	Sequence 4025, Ap

28	30.8	11.4	1875	4 US-09-252-991A-3919	Sequence 3919, Ap
29	30.6	11.3	897	4 US-09-902-540-4899	Sequence 4899, Ap
30	30.6	11.3	28493	4 US-09-902-540-1241	Sequence 1241, Ap
31	30.2	11.2	40181	4 US-09-949-016-17016	Sequence 17016, A
32	29.8	11.0	1056	3 US-09-067-782A-1	Sequence 1, Appl
33	29.6	11.0	26930	4 US-09-902-540-1328	Sequence 1328, Ap
34	29.4	10.9	488	4 US-09-621-976-2783	Sequence 2783, Ap
35	29.2	10.8	1101	4 US-09-205-258-119	Sequence 119, App
36	29	10.7	790	3 US-08-998-416-479	Sequence 479, App
37	29	10.7	2194	4 US-09-270-767-10411	Sequence 10411, A
38	29	10.7	2205	4 US-09-252-991A-11217	Sequence 11217, A
39	29	10.7	2454	4 US-09-252-991A-11038	Sequence 11038, A
40	29	10.7	13489	4 US-09-949-016-15911	Sequence 15911, A
41	29	10.7	20235	1 US-07-642-734C-3	Sequence 3, Appl
42	29	10.7	20235	3 US-08-439-009A-3	Sequence 3, Appl
43	28.8	10.7	439	3 US-09-067-782A-3	Sequence 3, Appl
44	28.8	10.7	777	4 US-09-902-540-3367	Sequence 3367, Ap
45	28.8	10.7	1383	1 US-08-484-044-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-09-949-001-30
Sequence 30, Application US/09949001
Patent No. 6825336
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO00789
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 30
LENGTH: 9173
TYPE: DNA
ORGANISM: Human
US-09-949-001-30

Query Match      22.5%; Score 60.8; DB 4; Length 9173;
Best Local Similarity 66.5%; Pred. No. 3.7e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY      53 CACAGCAAAACAACTATTGTAACAGCGGAGATCTAGACGCTGCCCTGACAAATCAT 112
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1877 CACTAGCAAAACAACTATTGTAACACTAGCTCTAGCGTGGCGGCTGCCAATCAT 1936

QY      113 TAACCC-----GTGCTGCCAGCAGCGCCCTTCAAGAGCCCTGGATATGCGCAGC 162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1937 TAACCTCTGTGTCAGAGTGGCGGCTGTGCTCTTTAAGATGCGGCTGTGTCAGC 1996

QY      163 CAGCATGTGTCACCTGCCCGCGAGACACAAACCCAGCAGCATTTGACACTGC-ACACGG 221
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1997 GAGATCGGCGCACCGCATTC-----CATTCAGCAGCATTCGCCGCGCGCGCG 2048

QY      222 CCATCTGCCAGAGAGCTGTGACCACTTC 253
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      2049 CCACCTCTCCAGAGAGCAGCTGCCACCGCTCC 2080

RESULT 2
US-09-949-001-36
Sequence 36, Application US/09949001
Patent No. 6825336
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF

```

FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match
Best Local Similarity 66.5%; Pred. No. 3,7e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 53 CACAGCAAAACAACTTATTTTGAACACGGGAGCTTGAACAGCTGCGCTGACATCAT 112
DB 1877 CACTAGCAAAACAACTTATTTTGAACACTGAGCTTGAACGCGCGCTGCAATCAT 1936

QY 113 TAACC-----GTGCTGCCGAGCAGCCCTTCAATAGCCCTGCGTATAGCCAGC 162
DB 1937 TAACCTCTGTGCAAGTGGCGGCGCTGTGCTTATAGGTGCGCGCTGTCTCAGC 1996

QY 163 GAGCATGTCTCACTGCGCGCGAGACACAACCGAGCGCATTTGAACACTGC-ACACGG 221
DB 1997 GAGCATGCGCGCACCGCATTC-----CATGCGAGCATTCGCGCGCGCGCGCG 2048

QY 222 CCATCTGCCAGAGAGCTGTGACCACTTC 253
DB 2049 CCACCTCCAGAGAGACTGCGCACCGCTCC 2080

RESULT 3
US-08-945-140-1
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Wallisop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816

TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match
Best Local Similarity 100.0%; Pred. No. 1,4e-07;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGGCGACGGGCACTCCGGTTCCTCGACTCGGCCCGAGTGA 51
DB 11 CATGGGCGACGGGCACTCCGGTTCCTCGACTCGGCCCGAGTGA 61

RESULT 4
US-08-791-849A-14
Sequence 14, Application US/08791849A
Patent No. 591449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,

```

; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
; FEATURE:
; NAME/KEY: Intron
; LOCATION: join(3219..3765, 3949..5916, 6009..6151,
; LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
; LOCATION: 9480..10162)
US-08-791-849A-14

Query Match          18.9%; Score 51; DB 2; Length 13011;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGGCGCACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 51
DB 3021 CATGGCGCACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 3071

RESULT 5
US-09-949-016-11786
; Sequence 11786, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11786
; LENGTH: 15108
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11786

Query Match          13.6%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.096;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 51
DB 2357 CACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 2399

RESULT 6
US-09-949-016-17205
; Sequence 17205, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17205
; LENGTH: 15108
; TYPE: DNA

; ORGANISM: Human
US-09-949-016-17205

Query Match          13.6%; Score 36.6; DB 4; Length 15108;
Best Local Similarity 90.7%; Pred. No. 0.096;
Matches 39; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 CACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 51
DB 2357 CACGGGCACTCCCGTGGTTCCTGACTCTGAGCCCGCCAGTGA 2399

RESULT 7
US-09-949-016-19194/c
; Sequence 19194, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19194
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19194

Query Match          13.0%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.075;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 130 CCAGCCCTTCATTAAGCCCTGGGTATGGCCAGCCAGATGTCACCTGCCCGAGACA 189
DB 478 CCAGCCCTTCCTGGGCGCCTGTCTCTTCATTCACGTTTAGCACCCTGCAGGTTAGGCC 419

QY 190 CAATCCAGCGAGCATTTGAACATGACACGCGCCATCTGCCAGAGACTGTGACCACCA 249
DB 418 CCCATCCAGAGACATCATCACCCGATATCAGACCCCTGCACCACTCATGATAGCC 359

QY 250 CTT 252
DB 358 CTT 356

RESULT 8
US-09-949-016-73784/c
; Sequence 73784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```


GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.

GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF cDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,942
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 03-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600

```

1 TELEFAX: (415) 543-5043
2 INFORMATION FOR SEQ ID NO: 3:
3 SEQUENCE CHARACTERISTICS:
4 LENGTH: 4447 base pairs
5 TYPE: nucleic acid
6 STRANDEDNESS: single
7 TOPOLOGY: linear
8 MOLECULE TYPE: cDNA
9 FEATURE:
10 NAME/KEY: CDS
11 LOCATION: 88..3162
12 FEATURE:
13 NAME/KEY: misc feature
14 LOCATION: 1..4447
15 OTHER INFORMATION: /product= "Pig DPP"
16
17 OS-08-991-944-3

```

Query Match	11.9%;	Score 32.2;	DB 3;	Length 4447;
Best Local Similarity	49.7%;	Pred. No. 1.8;		
Matches	82;	Conservative	0;	Mismatches 83;
			Indels	0;
			Gaps	0

Qy 96 GCTCCCTGCAATCATTAACCCGCGTGCAGCCAGCCCTTCAATGAAGCGCCGGTAT 155
Db 2275 GATGCTTTACAGCCACCAACCGCTCAGGCTCATGGATTAAAGCCGATGCACG 2333

0y GGCGGACGCGATGGTCCACTGCGCGCGAGACAAACCAGCAGCATTTGAACATCG 215
156 |||||
2335 CCTCGGCGGCGGTGGTGTGCGCAAGCGGACTACATACCGAGGAGCTGTTCGCGACGGCC 2394
Db |||||

QY 216 ACACGGCATCTGCCAGAGACTGTGACCACCACTTCCGGTACT 260
| | | | |
Db 2395 ATCAGACCAATTGCTTTGAGAGCTGTGACCACCACTTGTCGTGCT 2430

```

RESULT 15
US-09-949-016-11745
? Sequence 11745, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C0001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 11745
? LENGTH: 36180
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-11745

```

Query Match	11.9%;	Score 32;	DB 4;	Length 36180;
Best Local Similarity	53.1%;	Pred. No. 5.5;		
Matches 68;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;

Qy 105 ACATCATTAACCCGTGCTGCCGAGCCAGCCCTTCATATAGCCCTGGGATATGCCAGCCA 164
Db 2652 ACCCCCTTCCCCCGCGCGAGAACACAGAGCTAGTATGGCCCTGTGGGTTGGGCCCCCTA 2711

QY 165 GCATGTCACCTGCGCCCGGAGAACAAACCCAGCGAGATTGAACATCTGCACACGGCCA 224

Db 2712 GAGGGTTCACCTCGAGGCCAAGCCATGGCCACTGCGCCCGAGGGGGAATCCCTTGTTC 2711

225 TCTGCCCA 232

Db 2772 TCCGCCCA 2779

```
Search completed: September 1, 2005, 07:07:09
Job time : 87.4384 secs
```


Db 993 CCAGCATGCTCCAGCTGCCCGCCGAGACACAAACCCAGCAGCATGTGAACACTGCACACGG 992
Oy 273 CCATCTGCCAGAGAGCTGTGACCACTTCGGCTACTAGCTA 316
Db 993 CCATCTGCCAGAGAGCTGTGACCACTTCGGCTACTAGCTA 1036

RESULT 2
AC136382
LOCUS
DEFINITION
AC136382 185148 bp DNA linear HTG 01-NOV-2002
Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
*** 63 unordered pieces.
AC136382
AC136382.1 GI:24462257
HTG: HTGS PHASE1.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
AUTHORS
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D.,
Anyalelechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregiorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, V., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kows, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, Y., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorensbawa, L., Loulsegod, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhinney, S., McLeod, M., McNeill, T., Meenen, B., Milosavljevic, A.,
Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K.,
Morris, S., Muidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D.,
Newton, N., Nguyen, N., Norris, S., Nwakoeleneh, O., Okunolu, G.,
Olanunsgoon, A., Pal, S., Parker, K., Pasternak, S., Paul, H.,
Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polidexter, A.,
Popovic, D., Primus, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E.,
Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y.,
Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojias, A.,
Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S.,
Scott, G., Shatsman, S., Shen, H., Shetty, U., Shvartsbeyn, A.,
Sisson, I., Slater, C. D., Smales, D., Sneed, A., Sodergren, E.,
Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A.,
Sytek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S.,
Tingey, A., Trejos, Z., Uemami, K., Valdes, R., Vera, V., Villanasa, D.,
Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J.,
Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R.,
Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S.,
Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X.,
Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R.,
Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
REFERENCES
AUTHORS
Unpublished
2 (bases 1 to 185148)
Rat Genome Sequencing Consortium.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (01-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KOGS
Center clone name: CH230-97018
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990229
Consensus quality: 135613 bases at least Q40
Consensus quality: 140849 bases at least Q30
Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a working draft sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1011 1010: contig of 1010 bp in length
1111 1110: gap of unknown length
2609 2608: contig of 1498 bp in length
2709 2708: gap of unknown length
3915 3915: contig of 1207 bp in length
4015 4015: gap of unknown length
4016 5547: contig of 1532 bp in length
5548 5547: gap of unknown length
5548 7119: contig of 1472 bp in length
7120 7219: gap of unknown length
7220 8885: contig of 1666 bp in length
8886 8985: gap of unknown length
10259 10258: contig of 1273 bp in length
10259 10358: gap of unknown length
10359 11626: contig of 1268 bp in length
11627 11726: gap of unknown length
11727 13488: contig of 1762 bp in length
13489 13588: gap of unknown length
13589 15122: contig of 1534 bp in length
15123 15222: gap of unknown length
15223 16946: contig of 1724 bp in length
16947 17046: gap of unknown length
17047 18164: contig of 1118 bp in length
18165 18264: gap of unknown length
18265 19678: contig of 1414 bp in length
19679 19778: gap of unknown length
19779 21687: contig of 1909 bp in length
21688 21787: gap of unknown length
21788 22892: contig of 1105 bp in length
22893 22992: gap of unknown length
22993 24336: contig of 1344 bp in length
24337 24436: gap of unknown length
24437 26617: contig of 2181 bp in length
26618 26717: gap of unknown length
26718 28577: contig of 1860 bp in length
28578 28677: gap of unknown length
28679 30823: contig of 2146 bp in length
30824 30923: gap of unknown length
30924 32266: contig of 1343 bp in length
32267 32365: gap of unknown length
32367 33894: contig of 1528 bp in length
33895 33995: gap of unknown length
33996 35373: contig of 1379 bp in length

```
* 35374 35473: gap of unknown length
* 35474 37295: contig of 1822 bp in length
* 37296 37395: gap of unknown length
* 37396 38420: contig of 1025 bp in length
* 38421 38520: gap of unknown length
* 38521 39974: contig of 1454 bp in length
* 39975 40074: gap of unknown length
* 40075 42283: contig of 2209 bp in length
* 42284 42383: gap of unknown length
* 42384 43721: contig of 1338 bp in length
* 43722 43822: gap of unknown length
* 43822 46538: contig of 2717 bp in length
* 46539 46639: gap of unknown length
* 46639 48621: contig of 1983 bp in length
* 48622 48721: gap of unknown length
* 48722 50602: contig of 1881 bp in length
* 50603 50702: gap of unknown length
* 50703 53841: contig of 3139 bp in length
* 53842 53942: gap of unknown length
* 53942 56195: contig of 2254 bp in length
* 56196 56295: gap of unknown length
* 56296 58949: contig of 2654 bp in length
* 58950 59050: gap of unknown length
* 59050 61032: contig of 1983 bp in length
* 61033 61133: gap of unknown length
* 61133 62942: contig of 1810 bp in length
* 62943 63043: gap of unknown length
* 63043 66045: contig of 3003 bp in length
* 66046 66145: gap of unknown length
* 66146 68197: contig of 2052 bp in length
* 68198 68297: gap of unknown length
* 68298 72235: contig of 3938 bp in length
* 72236 72335: gap of unknown length
* 72336 75803: contig of 3468 bp in length
* 75804 78355: gap of unknown length
* 78356 78455: gap of unknown length
* 78456 81458: contig of 3003 bp in length
* 81459 81559: gap of unknown length
* 81559 84673: contig of 3115 bp in length
* 84674 84773: gap of unknown length
* 84774 87739: contig of 2966 bp in length
* 87740 87839: gap of unknown length
* 87840 91535: contig of 3656 bp in length
* 91536 91635: gap of unknown length
* 91636 95522: contig of 3887 bp in length
* 95523 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98831: gap of unknown length
* 98831 101618: contig of 2768 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117630: contig of 3966 bp in length
* 117631 117730: gap of unknown length
* 117731 122502: contig of 4772 bp in length
* 122503 122602: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126645 126744: gap of unknown length
```

Query Match 66.8%; Score 214.4; DB 2; Length 185148;
Best Local Similarity 97.3%; Pred. No. 6.1e-53;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 93 TCGCCCAATCTCAACAGCAAACTTATTTTGAACAGCGGGATCTTAGACCGCTGCC 152
DB 99243 TGTGGAGACTCACAAGCAAACTTATTTTGAACAGCGGGATCTTAGACCGCTGCC 99302
QY 153 CTGACAATCATTAACCCGCTGCGAGCGACCCCTTCAATAGCGCCCTGGGATAGCCAG 212

```
DB 99303 CTGACAATCATTAACCCGCTGCGAGCGACCCCTTCAATAGCGCCCTGGGATAGCCAG 99362
QY 213 CCAGCATGATCTCACTCCCGCGAGACAAACCCAGAGAGATTGAACACTGCACGCG 272
DB 99363 CCAGCATGATCTCACTCCCGCGAGACAAACCCAGAGAGATTGAACACTGCACGCG 99422
QY 273 CCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
DB 99423 CCATCTGCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 99466

RESULT 3
MILIGF 1363 bp DNA linear ROD 01-AUG-1996
LOCUS M.musculus gene for insulin-like growth factor binding protein-1.
DEFINITION X67493
ACCESSION X67493.1 GI:52699
VERSION insulin-like growth factor binding protein-1.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1363)
AUTHORS Mohn,K.L., Bucan,M., and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression during liver regeneration
JOURNAL Hepatology 19 (3), 656-665 (1994)
MEDLINE 94164648
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohn,K.L., Weddle,J.R., and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential insulin-responsive sequence and conservation of all intron/exon boundaries
JOURNAL Nucleic Acids Res.
REFERENCE 3 (bases 1 to 1363)
AUTHORS Taub,R.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard Hughes Medical Institute, Clinical Research Bldg., Room 475, 422 Curie Boulevard, Philadelphia, PA 19104-6145, USA

FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="Genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="JW2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus sequences"
764..768
/note="exon"
792..1336
/number=1
792
/note="rRNA cap site"
964..>1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"
```

/translation="MPEFLTVSWPFLILLISFOIGVAGAPQPMHCAPCTABRLGCP
PVPASCPETSRPAGCGCCCTCTCALPMGACAGVATARCACQGLSRALPGEPRPLHALTRG
QGASLPBPAPAPARTSLFSSQNH"

ORIGIN

Query Match 58.8%; Score 188.8; DB 10; Length 1363;
Best Local Similarity 90.2%; Pred. No. 2.1e-45;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 93 TGGGCCCATGTGCACAGCAAAACAACCTATTGTAACAACGGGATCCTAGACGCTGCC 152
|||
DB 667 TGTGTAGAGCTCACAGCAAAACAACCTATTGTAACAACGCGGCTCTAGACGCTGCC 726
|||
QY 153 CTGACATCATTTAAACCCCTGTGCTGCCGAGCAGCCCTTCATTAAGGCTGGTATGGCCAG 212
|||
DB 727 CTGACATCATTTAACTGTGTGCCGACAGCAGCCCTTCATTAAGGCTGTGGTATGACACGAG 786
|||
QY 213 CCAGCATGTGTCACTGCCCCCGGAGACACAAACCCAGCAGCATTTGAACACTGCACACGG 272
|||
DB 787 CCAGCATGTGTCACTGCCCCCGGAGACACACCCAGCAGCATTTGAACACTGCACACGG 846
|||
QY 273 CCATCTGCCAGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 316
|||
DB 847 CCGTCTGCCAGAGAGCTGTGACCAACCACTCCGCTACTACTA 890
|||

RESULT 4
AL607124
LOCUS Mouse DNA sequence from clone RP23-20C9 on chromosome 11, complete
DEFINITION
ACCESSION AL607124
VERSION AL607124.15 GI:20145926
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Oliver, K.
Direct Submission
Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humgen@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:
SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.choi.org/bacpac/home.htm
VECTOR: pBACe3.6.

FEATURES

source

Location/Qualifiers
1..192843
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

/clone="RP23-20C9"
/clone_1ib="RPI-23"

ORIGIN

Query Match 58.8%; Score 188.8; DB 10; Length 192843;
Best Local Similarity 90.2%; Pred. No. 2.8e-45;
Matches 202; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 93 TGGGCCCATGTGCACAGCAAAACAACCTATTGTAACAACGGGATCCTAGACGCTGCC 152
|||
DB 2419 TGTGTAGAGCTCACAGCAAAACAACCTATTGTAACAACGCGGCTCTAGACGCTGCC 2478
|||
QY 153 CTGACATCATTTAAACCCCTGTGCTGCCGAGCAGCCCTTCATTAAGGCTGGTATGGCCAG 212
|||
DB 2479 CTGACATCATTTAACTGTGTGCCGACAGCAGCCCTTCATTAAGGCTGTGGTATGACACGAG 2538
|||
QY 213 CCAGCATGTGTCACTGCCCCCGGAGACACAAACCCAGCAGCATTTGAACACTGCACACGG 272
|||
DB 2539 CCAGCATGTGTCACTGCCCCCGGAGACACACCCAGCAGCATTTGAACACTGCACACGG 2598
|||
QY 273 CCATCTGCCAGAGAGCTGTGACCAACCACTCCGCTACTAGCTA 316
|||
DB 2599 CCGTCTGCCAGAGAGCTGTGACCAACCACTCCGCTACTACTA 2642
|||

RESULT 5
BC013345
LOCUS Mus musculus insulin-like growth factor binding protein 1, mRNA
DEFINITION (CDNA clone MGC:14075 IMAGE:4161889), complete cds.
ACCESSION BC013345
VERSION BC013345.1 GI:15426482
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1569)
Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G.,
Klausner, R. D., Collins, F. S., Wagner, L., Shennan, C. M., Schuler, G. D.,
Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K.,
Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Heide, F.,
Dichtenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L.,
Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L.,
Schaefer, T. E., Brownstein, M. J., Uedl, T. B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J.,
Abramson, R. D., Mullan, S. J., Bosak, S. A., McEwan, P. J.,
McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S.,
Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hui, S. W.,
Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A.,
Fahey, J., Heltan, E., Kettman, M., Madan, A., Young, A. C., Shevchenko, Y.,
Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y.,
Bonfard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D.,
Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M.,
Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smalins, D. E.,
Scherer, A., Schein, J. E., Jones, S. J., and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (31-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

COMMENT

NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

Contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILN.ac: http://image.llnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1. 1569

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="MGC:14075 IMAGE:416189"

/issue_type="Liver, normal. 5 month old male mouse."

/clone_id="NCI CGAP_L19"

/lab_host="DH10B"

/note="Vector: PCMV-SPORT6"

1. 1569

/gene="Igfbp1"

/note="synonym: IGFBP-1"

/db_xref="LOCUSID:16006"

/db_xref="MGI:36436"

204. 1022

/gene="Igfbp1"

/codon_start=1

/product="insulin-like growth factor binding protein 1"

/protein_id="AAH13345.1"

/db_xref="GI:15426483"

/db_xref="MGI:16006"

/db_xref="MGI:36436"

/translation="MPELVTVSWPELILSFQIGVAGAPQPMHCAPICTARLGLCP
PVPAPELSPRAGCCGCTCCALPMGACGATATACAGLSRAIPGEPRLHALTRG
QACVPEPATPTASGLSIEKEAKASVPRVPESEMTSEQLDSFHLMASSSD
OPILMVAISTYSMRAREIADIKKKEPCRELIVLERLAAQKAGEBIYKFLPN
CNKNGFYHSKOCETSLDGEARLQMCVYPWSGRIIPGSLERDPPCHQYFNVQN"

ORIGIN

Query Match 30.6%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 2.7e-18;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 206 TGGCCAGCAGCATGTGTCACATGCCCGCCGAGACACAAACCCAGCAGCATTTGAACACTG 265
|||
DB 20 TGGGACAGCAGCATGTGTCACATGCCCGCCGAGACACACCCAGCAGCATTTGAACACTG 79
|||
QY 266 CACACGGCCATTTGCCAGAGAGCTGTGACACCACTTCGGCTACTAGCTA 316
|||
DB 80 CACACGGCCGTCTGCCAGAGAGCTGTGACACCACTTCGCTACTATCTA 130
|||

RESULT 6 1507 bp mRNA linear ROD 22-MAR-2004
LOCUS AY560836
DEFINITION *Spermophilus tridecemlineatus* insulin-like growth factor binding
protein 1 (IGFBP1) mRNA, complete cds.
ACCESSION AY560836
VERSION AY560836.1 GI:45505308

SOURCE
KEYWORDS
ORGANISM
Spermophilus tridecemlineatus (thirteen-lined ground squirrel)
Spermophilus tridecemlineatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
Spermophilus.

REFERENCE
AUTHORS
TITLE
1 (bases 1 to 1507)
Li,Y., Klimans,D. and Hallenbeck,J.M.
Cloning and characterization of insulin-like growth factor binding
protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL

Unpublished

2 (bases 1 to 1507)

REFERENCE

Li,Y., Klimans,D. and Hallenbeck,J.M.

AUTHORS

Direct Submission

TITLE

Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
Bethesda, MD 20892, USA

JOURNAL

location/Qualifiers

FEATURES

1. 1507

/organism="Spermophilus tridecemlineatus"

/mol_type="mRNA"

/db_xref="taxon:43179"

1. 1507

/gene="IGFBP1"

689. 1507

/gene="IGFBP1"

/codon_start=1

/product="insulin-like growth factor binding protein 1"

/protein_id="AA567029.1"

/db_xref="GI:45505309"

/translation="MPEVPAAGLMPFLLAVQSTVASSTQPMHCAPICTARLGLCP
FVPSCEPLSPRAGCCGCTCCALPMGACGATATACAGLSRAIPGEPRLHALTRG
QACVPEPATPTASGLSIEKEAKASVPRVPESEMTSEQLDSFHLMASSSD
OPILMVAISTYSMRAREIADIKKKEPCRELIVLERLAAQKAGEBIYKFLPN
CNKNGFYHSKOCETSLDGEARLQMCVYPWSGRIIPGSLERDPPCHQYFNVQN"

ORIGIN

Query Match 24.7%; Score 79.2; DB 10; Length 1507;
Best Local Similarity 67.1%; Pred. No. 1.3e-12;
Matches 149; Conservative 0; Mismatches 58; Indels 15; Gaps 2;

QY 93 TGGCCCATGTGCAGAGAAACAACTATTTGAAACAGGGGATCCTAGACGCTGCC 152
|||
DB 420 TGGGAGAAAGAGAGAGAAACAACTATTTGAAACAGGGGATCCTAGACGCTGCCGCGG 479
|||
QY 153 CTGACAACTATTAA-----CCCGTGTGTCGCGAGCCAGCCCTTATAGGCGCTGG 202
|||
DB 480 CTGGCAATCTTAATTCTTCTAGTCCAACTGAGAGTGAACCGCGCCTTTATAGGAAGG 539
|||
QY 203 GTATGCGCAGCAGCATGTGCTGCTGCCCGCCGAGACAAACCCAGCAGCATTTGAACA 262
|||
DB 540 CTGCACCCAGCAGCATGTGCTGCTGCCCGCCGAGCAGTCAAGATTGACACTATCA-- 597
|||
QY 263 CTGCACAGGCCTATGTCGCCAGAGCTGTGACCACTTC 304
|||
DB 598 ---GCCACTGCGAGCTGCCAGAGAGCATTTGACCACTGTCC 636
|||

RESULT 7
LOCUS AX401932 1500 bp DNA linear PAT 06-JUN-2002
DEFINITION Sequence 1608 from Patent WO0210453.
ACCESSION AX401932
VERSION AX401932.1 GI:21338112

SOURCE
KEYWORDS
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
TITLE
1
Mendrick,D., Porter,M.W., Johnson,K.R., Caastle,A.L. and
Elschoff,M.R.

Molecular toxicology modeling
Patent: WO 0210453-A 1608 07-FEB-2002;
Gene Logic, Inc. (US)

FEATURES

location/Qualifiers

1. 1500

/organism="Rattus norvegicus"

/mol_type="unassigned DNA"

/db_xref="taxon:10116"

/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

```

Query Match          23.8%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
      |||||||
DB 8 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 67

QY 299 CACTTCGGCTACTACTA 316
      |||||||
DB 68 CACTTCGGCTACTACTA 85

RESULT 8
AX827271 1500 bp DNA linear PAT 12-DEC-2003
DEFINITION Sequence 5 from Patent EP1344834.
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
          F. HOFMANN-LA ROCHE AG (CH)
FEATURES
          source
            1. 1500
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          23.8%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
      |||||||
DB 8 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 67

QY 299 CACTTCGGCTACTACTA 316
      |||||||
DB 68 CACTTCGGCTACTACTA 85

RESULT 9
RATIGFB
LOCUS RATIGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (rIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
          Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 1500)
AUTHORS Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE 91141487
PUBMED 1705004
COMMENT source text: Rat, cDNA to mRNA.
FEATURES
          location/Qualifiers
            1. 1500
              /organism="Rattus norvegicus"
              source

```

```

/mol_type="mRNA"
/strain="fisher"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AA41380.1"
/db_xref="GI:204733"
/translation="MPEFLTVSNPFLILSFQVRVVAAGAPQPMHCAPCTARIELCP
PVPASCEISRDPAGCGCCPTCALPLAGAAVATRCAGLSRALPEPRPLALTRG
QGACVLEPAAPATSLISGSOHEFAKAAVASDELAESPEMTEEOILDSFHMARSRED
QPLIMNAISTYSMRARBITDCLKKEPCORELYKLEBRLAAAOOKAGDEIKFYFLPN
CNKGPFYSKOCETSLDGEAGLCMCVYPMGSKTIPSLFETRGDINCHQFNVQV"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match          23.8%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 8.9e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
      |||||||
DB 8 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 67

QY 299 CACTTCGGCTACTACTA 316
      |||||||
DB 68 CACTTCGGCTACTACTA 85

RESULT 10
AX163782 5001 bp DNA linear PAT 22-JUN-2001
LOCUS AX163782
DEFINITION Sequence 46 from Patent WO0138579.
ACCESSION AX163782
VERSION AX163782.1 GI:14544878
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
          Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1
AUTHORS Gould-Rothberg, B.E., DiPippo, V.A., Rameeh, T.M. and Gerwein, R.W.
TITLE Method of identifying toxic agents using msaid-induced differential
JOURNAL gene expression in liver
PUBMED Patent: WO 0138579-A 46 31-MAY-2001;
          Curagen Corporation (US)
FEATURES
          location/Qualifiers
            1. 5001
              /organism="Rattus norvegicus"
              /mol_type="unassigned DNA"
              /db_xref="taxon:10116"

ORIGIN
Query Match          23.8%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 9.5e-12;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 239 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 298
      |||||||
DB 1 CACAAACCCAGCGAGCATTTGAACACTGACACGGCCATCTGCCAGAGAGCTGTGACAC 60

QY 299 CACTTCGGCTACTACTA 316
      |||||||
DB 61 CACTTCGGCTACTACTA 78

RESULT 11
RATIGFBA

```


DB	I	CACAAACCAGCGAGCATTGAACACTGCACAGGCGATCTGGCCAGAGAGCTGTGACCAC	60
OY	299	CAC TTC CGCTACTAGCTA	316
DB	61	CAC TTC CGCTACTAGCTA	78
RESULT 12			
LOCUS	BC078889		
DEFINITION	BC078889	1510 bp	mRNA
ACCESSION	Rattus norvegicus insulin-1-like growth factor binding protein 1,		linear ROD 03-AUG-2004
VERSION	MRA (CDNA clone MGC:93595 IMAGE:7129185), complete cds.		
KEYWORDS	BC078889		
SOURCE	MGC.	GI:50927646	
ORGANISM	Rattus norvegicus (Norway rat)		
	Rattus norvegicus		
	Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1	(bases 1 to 1510)	

AUTHORS	Staubsberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, P. S., Wagner, L. H., Sherman, C. M., Schuler, G. D., Altschuler, S. F., Zeeberg, B., Bielow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Heitel, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Slepton, M., Soares, M. B., Bonaldi, M. F., Casavant, T. L., Schetz, T. E., Brownstein, M. J., Usdin, T. B., Tohyama, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hylk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bonifard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywicki, M. I., Skalske, U., Smailus, D. E., Schmechel, A., Schein, J. E., Jones, S. U., and Marz, M. A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCE	12477932
AUTHORS	2 (bases 1 to 1510)
TITLE	Director MGC Project.
JOURNAL	Direct Submission
COMMENT	Submitted (02-VUG-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
CONTACT	Contact: MGC help desk
EMAIL	Email: cgabs-remail.nih.gov
TISSUE	Tissue Procurement: Howard Jacobs
CNA	CNA Library Preparation: Expressed Genomics
CDNA	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA	DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
WEB	Web site: http://www-shgc.stanford.edu
CONTACT	Contact: (Dickson, Mark) mcd@paxil.stanford.edu
DICKSON	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Series: IRAP Plate: 184 Row: e Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6981079, location/Qualifiers
SOURCE	1. 1510 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116"

gene

```

/clone="MGC:93595 IMAGE:7129185"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_lib="NIH MGC_236"
/lab_host="DH10B"
/notes="Vector: pEXpress1"
1..1510
/gene="Igfbp1"
/notes="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="locusid:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
142..960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH78889.1"
/db_xref="GI:50927647"
/db_xref="LocusID:25685"
/db_xref="RATMAP:44422"
/db_xref="RGD:2872"
/translation="MPEPLTVSWPFLILSFQVRVVGAPQPMHCAPCTARLILCP
PYVASCPRISRAGCCCTCALPIGAAGVATACAGISCRALPGRPRPHALTREG
QACVDEPAAPATSSLSGSHBEAKAAVAEDELASPEMTEQULDSFHLMAERED
QPLMNAISTYSMPAREITDLKMKPEQRELYVLEBLAAQAKADELYKFLPN
CKNGFYHSHKOCETSLDGEAGLCWCVPWSGKKIPGSLERDPCNHOYFNVQN"

```

ORIGIN

Query Match 20.4%; Score 65.4; DB 10; Length 1510;
 Best Local Similarity 98.5%; Pred. No. 1.7e-08;
 Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 250 CGAGATTGAACAGTCACAGCGCCATCTGCCAGAGAGCTGTGACCACTTCGCGCTA 309
 |||||
 DB 1 CGAGATTGAACAGTCACAGCGCCATCTGCCAGAGAGCTGTGACCACTTCGCGCTA 60

QY 310 CTAGCTA 316
 |||||
 DB 61 CTATCTA 67

RESULT 13
 AY095345 3886 bp DNA linear PRI 30-DEC-2002
 LOCUS Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
 DEFINITION Gene, partial cde.
 ACCESSION AY095345
 VERSION AY095345.1 GI:20853764
 KEYWORDS
 SOURCE
 ORGANISM
 Papio anubis (olive baboon)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Papio.
 1 (bases 1 to 3886)
 Kim,J.J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
 Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
 Regulation of insulin-like growth factor binding protein-1 promoter
 activity by FKHR and HOKA10 in primate endometrial cells
 JOURNAL Biol. Reprod. 68 (1), 24-30 (2003)
 PUBMED 12493691
 2 (bases 1 to 3886)
 Kim,J.J., Jaffe,R.C. and Fazleabas,A.T.
 Direct Submission
 TITLE Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
 JOURNAL University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
 USA

FEATURES
 source
 1..3886
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /db_xref="taxon:9555"
 <3655..>3886
 /gene="IGFBP-1"

mRNA

```

<3655..>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655..>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAH23273.1"
/db_xref="GI:20853765"
/translation="MSEVPVAVVLLITVQYGVYASAPWQCAPCSAEKALCPV
PASCSEVTRSHAGCGCCPMCALPIGAAGVATAR"

```

CDS

ORIGIN

Query Match 19.9%; Score 64; DB 9; Length 3886;
 Best Local Similarity 60.6%; Pred. No. 4.8e-08;
 Matches 149; Conservative 0; Mismatches 80; Indels 17; Gaps 2;

QY 69 TCCAGGAACCAAGGAGTGCCTGCGCCCATGTCAACAAACAACTTATTTGA 128
 |||||
 DB 3331 TCTCCCAACCTGCGGTTTGCGTAGAGCCCTTGCGCGCACTAGCAAAACAACTTATTTGA 3390

QY 129 ACAGGGATCTTGACAGCTGCTGCTGACATCATTAACC-----GTGCTCCG 178
 |||||
 DB 3391 ACACCTCAGCTCTGCGCGCGCGCTGCAATCATTAACCTCTGTCAGAGTGCGCGG 3450

QY 179 AGCCAGCCCTTCATTAAGCCCTGCGGTATGCGCAGCAGATGTCCACTGCGCGCAGA 238
 |||||
 DB 3451 CCTGTGCTTTATTAAGCGCGCGCTGTGTCCAGAAAGCATGCGCCATCCCATC 3510

QY 239 CACAAACCCAGCAGCATTTGAACACTGCACACGCGCATCTGCCAGAGAGCTGACAC 298
 |||||
 DB 3511 CAGCAAGC-----ATCTGCGCGCGCGCGCGCGCCACCTCCAGAGAGCATGCGCAC 3563

QY 299 CACTTC 304
 |||||
 DB 3564 CGCTCC 3569

RESULT 14
 AX409747 6128 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 2394 from Patent W00229103.
 DEFINITION AX409747
 ACCESSION AX409747.1 GI:21442452
 VERSION AX409747.1 GI:21442452
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 Alvarez,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
 Gene expression profiles in liver cancer
 PATENT: WO 0229103-A 2394 11-APR-2002;
 JOURNAL GENE LOGIC INC (US)

FEATURES
 source
 1..6128
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="EMBL/Genbank Accession No. W4587"

ORIGIN

Query Match 19.9%; Score 63.8; DB 6; Length 6128;
 Best Local Similarity 63.2%; Pred. No. 5.6e-08;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 69 TCCAGGAACCAAGGAGTGCCTGCGCCCATGTCAACAAACAACTTATTTGA 128
 |||||
 DB 438 TCTCCCAACCTGCGGTTTGCGTAGAGCCCTTGCGCGCACTAGCAAAACAACTTATTTGA 497

QY 129 ACAGGGATCTTGACAGCTGCTGCTGACATCATTAACC-----GTGCTCCG 178
 |||||
 DB 498 ACACCTCAGCTCTGCGCGCGCGCTGCAATCATTAACCTCTGTCAGAGTGCGCGG 557

QY	179	AGCCGACCCCTTATAGAGCCCTGGGTATGGCCAGCAGCATGATGTCACCTGGCCGCCGAGA	238
Db	558	CCTGGCCCTTATATAGGTGCGCGCTGTGTCCAGGAGCATCGCCACCGCATTC----	613
QY	239	CACAAACCCAGGAGCATTTGAACACTGC-ACACGGCCATCTGCCAGAGAGCTGTGACCA	297
Db	614	----CATCCAGGAGACATCTGCCGCGCGCGCCGCCACCCCTCCAGAGAGCATGCGCA	669
QY	298	CCACTTC	304
Db	670	CCGCTCC	676
RESULT 15			
HUMIGFBP1A			
LOCUS		6128 bp	DNA
DEFINITION		Human insulin-like growth factor binding protein (hIGFBP1) gene,	linear
ACCESSION		complete cds.	FRI 08-NOV-1994
VERSION		M74587	
KEYWORDS		insulin-like growth factor binding protein.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
		1 (bases 1 to 6128)	
TITLE		Ehrenborg, E., Larsson, C., Stern, I., Jansson, M., Powell, D.R. and	
		Luthman, H.	
		Contiguous localization of the genes encoding human insulin-like	
		growth factor binding proteins 1 (IGBP1) and 3 (IGBP3) on	
		chromosome 7	
JOURNAL		Genomics 12 (3), 497-502 (1992)	
MEDLINE		92217971	
PUBMED		1373120	
COMMENT		Original	
FEATURES		source text: Homo sapiens (tissue library: lambda	
		Charon4A) liver DNA.	
source		location/Qualifiers	
		1..6128	
		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		/map="7p13-p12"	
		/feature_type="liver"	
		/feature_1fb="lambda Charon4A"	
		525..529	
		/gene="IGFBP1"	
		/note="G00-120-075"	
		569..573	
		/gene="IGFBP1"	
		/note="G00-120-075"	
		join(597..1110,2657..2826,4041..4169,5069..5769)	
		/gene="IGFBP1"	
		join(597..1110,2657..2826,4041..4169,5069..5769)	
		/gene="IGFBP1"	
		/product="insulin-like growth factor binding protein 1"	
		/note="G00-120-075"	
		597..1110	
		/gene="IGFBP1"	
		/note="G00-120-075"	
		join(762..1110,2657..2826,4041..4169,5069..5200)	
		/gene="IGFBP1"	
		/codon_start=1	
		/product="insulin-like growth factor binding protein 1"	
		/protein_id="AA52784.1"	
		/db_xref="GI:184812"	
		/db_xref="GDB:G00-120-075"	
		/translation="MSEVPAVAVWLIVLLLVQVVTAGAPWQCAPCSAEKALACPVP	
		SASSESDVASGCCCPWCAPLGAAGVATARCAGISGALGEGOOPIHALITRGGG	
		ACVQESDVSAPHAABAGSPSESTETITETELINFLMAPSEBDHSILMDAISTYG	
		SKALHVNIIKKWKEPCRIETLYRVVSLAKQSTGSEISTSKYLPNCNKGFIYSRQCE	
		TSNMGEGALCKCVTFPMNGKRIPLGSPETIRGDENCMYFNVO"	
		762..836	
s1g_peptide			

```

mat_peptide      /gene="IGFBP1"
                  /note="G00-120-075"
                  join(837..1110,2657..2826,4041..4169,5069..5197)
                  /gene="IGFBP1"
intron           /product="insulin-like growth factor binding protein 1"
                  /note="G00-120-075"
                  1111..2656
                  /gene="IGFBP1"
variation        /note="G00-120-075"
                  1405
exon             /gene="IGFBP1"
                  /note="G00-120-075"
                  2657..2826
                  /gene="IGFBP1"
intron          /note="G00-120-075"
                  2827..4040
                  /gene="IGFBP1"
exon            /note="G00-120-075"
                  4041..4169
                  /gene="IGFBP1"
intron         /note="G00-120-075"
                  4170..5068
                  /gene="IGFBP1"
variation      /note="G00-120-075"
                  4736
                  /gene="IGFBP1"
variation      /note="G00-120-075"
                  4780
                  /gene="IGFBP1"
variation      /note="G00-120-075"
                  4820
                  /gene="IGFBP1"
exon           /note="G00-120-075"
                  5069..5769
                  /gene="IGFBP1"
                  /note="G00-120-075"

ORIGIN
Query Match      19.9%; Score 63.8; DB 9; Length 6128;
Basic Local Similarity 63.2%; Pred.No. 5.6e-08;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3

QY    TCCAGGAACCA CGGGAGTGC CCGCTGTG CCGCATGTCA CAAGAACA AACTTTAT TTGA   128
      |||       |||       |||       |||       |||       |||
DB    TCCTCCCA CCA CGSGTTTG CTGTAAGG CCTTGG GTGCACTA GCACAA AACAACTTA TTTTGA   497

QY    129 ACACGGGAGT CTTAGC AGCTGTG CCGCTTGA CAAATTCA TTAAACC-----GTGCTGCCG 178
      |||       |||       |||       |||       |||       |||
DB    498 ACACTCA GCTCTTA GCGTGG CCGCTGTG CAAATCA TTAACTTC TGTCAGT GAAGTGG CCGCG 557

QY    179 AGCCAGC CTCCTCA TTAAGCC CTGTGG GTATGG CCAAGC CAGATGTG TCCA CTGCGCCGCC GAGA 238
      |||       |||       |||       |||       |||       |||
DB    558 CCTGTGC CTTTA TAAGGTG CGCCTGTG TCCAGC GAGCATCG GCACCCGCCATCC---- 613

QY    239 CACAACCC ACGAGCA TTGAAC ACTGC -ACACGCC CATTGT GCCCAG AGAGCTGT GACCA 297
      |||       |||       |||       |||       |||       |||
DB    614 ----CA TCCAGC GAGCAT TGCAGC CGCGCC GCGCCG CCACTCT CCAGAG AACTGTG CCA 669

QY    298 CCACTTC 304
      |||       |||
DB    670 CCGCTCC 676

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 2440.45 Seconds
(without alignment)
5006.706 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321
Sequence: 1 taccactggggccagagatcc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hlc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	30.6	1013	2	BF236974
2	87.8	27.4	310	6	BY783538
3	87.8	27.4	331	6	BY794229
4	87.8	27.4	354	6	BY771317
5	86.8	27.0	401	1	A1785818
6	86.8	27.0	480	1	A1196314
7	86.8	27.0	706	1	A1530146
8	86.8	27.0	765	1	A1098594
9	86.8	27.0	785	1	A1530313
10	86.8	27.0	799	1	A1529939
11	86.8	27.0	811	6	CA478518
12	86.8	27.0	848	1	A1790802
13	86.8	27.0	852	1	A1528304
14	85.2	26.5	305	1	A1196154
15	82.2	25.6	680	1	A1785039
16	81.4	25.4	846	7	CO573026
17	80.4	25.0	692	7	CV127049
18	79.4	24.7	713	7	CO560662
19	75.4	23.5	748	7	CO575629
20	73.8	23.0	694	7	CV117001
21	72.4	22.8	632	7	W30013
22	72.4	22.6	615	7	BB660958
23	72.4	22.6	618	6	CD561711
24	72.4	22.6	632	1	A1892189

25	67.4	21.0	488	1	AA674302
26	65.4	20.4	801	7	CK472246
27	64.4	20.1	840	7	CK473709
28	62.8	19.6	759	1	AA105355
29	61.8	19.3	269	2	BB604790
30	61.6	19.2	587	2	AW916237
31	60.8	18.9	1589	3	CR621807
32	60.8	18.9	1601	3	CR95377
33	59	18.4	545	1	AA060360
34	47.2	14.7	659	8	AZ840793
35	38.8	12.1	698	5	BU684425
36	36.8	11.5	428	2	BF829782
37	36.8	11.5	549	7	CV014020
38	36.8	11.5	849	9	CNS0208E
39	36.6	11.4	370	7	CN402790
40	36.6	11.4	459	7	CN402791
41	36.6	11.4	466	7	CN402792
42	36.6	11.4	497	6	CD635411
43	36.6	11.4	498	6	CD635412
44	36.6	11.4	744	7	CN402794
45	36.6	11.4	771	5	BX360059

ALIGNMENTS

RESULT 1
BF236974 1013 bp mRNA linear EST 14-NOV-2000
LOCUS 60202605F1_NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:416189 5',
DEFINITION mRNA sequence.

ACCESSION BF236974 GI:11150891
VERSION BF236974.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-bases 1 to 1013
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov

COMMENT

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence stop: 581.
Location/Qualifiers
1..1013

FEATURES

source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:416189"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6, site 1: NotI; site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 30.6%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 2.6e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 206 TGGCAGCAGCAGTGTCTACTGCCCGCCGAGACACAAACGAGCAGATTGAACACTG 265

Db	9	TGGGAGGCCAGCATGTCTCCACTTCGCCCGGAACACACACCAGCAGATTGAACACTTG
Oy	266	CACACGGCCATCTGCCGCAGAGAGCTGTGTACCAACCACTTCCGCTACTAGCTTA 316 CACACGGCGCTCTGCCGCAGAGAGCTGTGTACCAACCACTTCCGCTACTAGCTTA 319
Db	69	
	RESULT 2	
	BY783538	310 bp mRNA linear EST 23-MAR-2004
	LOCUS	BY783538 RIKEN full-length enriched, 17.5 days embryo whole body
	DEFINITION	Mus musculus cDNA clone U930176D05 5', mRNA sequence.
	ACCESSION	BY783538
	VERSION	BY783538.1 GI:39710177
	KEYWORDS	EST.
	SOURCE	Mus musculus (house mouse)
	ORGANISM	Mus musculus
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 310) Carinici, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Otsu, N., Fukuda, S., Sato, K., Warahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustinich, S., Beisel, K., Pavan, W., Aldrich, V., Nakagawara, A., Held, W. A., Iwata, H., Konno, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, J. M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mommaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y. Targeting a complex transposome: the construction of the mouse full-length cDNA encyclopedia Genome Res. 13 (6B), 1273-1289 (2003)
	TITLE	Contact: Yoshihide Hayashizaki
	JOURNAL	Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
	MEDLINE	The Institute of Physical and Chemical Research (RIKEN)
	PUBMED	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.
	COMMENT	Location/Qualifiers 1..310 /organism="Mus musculus" /mol_type="mRNA" /strain="CS7BL/6J" /db_xref="taxon:10090" /clone="U930176D05" /tissue.type="whole body" /dev_stage="17.5 days embryo" /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"
FEATURES	source	
ORIGIN	Query Match	27.4% Score 87.8; DB 6; Length 310;
	Best local Similarity	92.9%; Pred. No. 2.1e+14;
	Matches	92; Conservative 0; Mismatches 7; Indels 0; Gaps 0
Oy	218	ATGTCACACTGCGCCGCGAGACACAACCCAGCGAGATTGAACACTGCACAGGCCATTC 277
Db	2	ATGTCACACTGCGCCGCGAGACACAACCCAGCGAGATTGAACACTGCACAGGCCATTC 61
Oy	278	TGCCAGAGAGCTGTACCACTTCGCGCTACTAGCTTA 316

Db	62	TGCCAGAGACTGTGACCACTATTGCACCTACTATCTA	100
RESULT 3			
LOCUS	BV794229	331 bp	mRNA linear EST 23-MAR-2004
DEFINITION	BV794229 RIKEN full-length enriched 17.5 days embryo whole body		
ACCESSION	BV794229		
VERSION	BV794229.1	GI:39720868	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus. 1 (bases 1 to 331)		
AUTHORS	Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanishi, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusabe, M., Gustincich, S., Beisel, K., Pavan, W., Altinli, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, T. K., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawaji, J. and Hayashizaki, Y.		
TITLE	Targeting a complex transcripome: the construction of the mouse full-length cDNA encyclopedea		
JOURNAL	Genome Res. 13 (6B), 1273-1289 (2003)		
MEDLINE	22703353		
PUBMED	12819125		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/ cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (http://genome.gsc.riken.jp/) for further details.		
FEATURES	source	location/Qualifiers	
	1..331		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="L930292H24"		
	/tissue_type="whole body"		
	/dev_stage="17.5 days embryo"		
	/clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"		
ORIGIN			
Query Match	27.4%	Score 87.8;	DB 6;
Best local Similarity	92.9%	Pred. No. 2.1e-14;	
Matches	92;	Conservative 0;	Mismatched 7;
		Indels 0;	Gaps 0;
Oy	218	ATGGTCACCTCGCCGCGAGACAATAAACCCGAGCATTTGAACATGTGCACACGGCCATC	277
Db	2	ATGGTCCACTGCCCGGAGACACACACCACGACGACATTTGAACATGTGCACACGGCCGTC	61
Oy	278	TGCCAGAGACTGTGACCACTTCGCTACTATAGCTA	316
Db	62	TGCCAGAGACTGTGACCACTATTGCACCTACTATCTA	100
RESULT 4			
	BV771317		

VERSION	AI785818.1	GI:5335354
KEYWORDS	EST.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 401)	
AUTHORS	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Smaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E., Rohm,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Marston,R. and Wilson,R.	
TITLE	The WashU-NCI Mouse EST Project 1999	
JOURNAL	Unpublished (1999)	
COMMENT	Other_ESTs: u178h05.xl Contact: Marra M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LINDL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:972993	
FEATURES	Seq primer: custom primer used High quality sequence stop: 126.	
SOURCE	Location/Qualifiers	
	1..401	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="C57BL"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:1888569"	
	/sex="female"	
	/dev_stage="adult"	
	/lab_host="DH10B"	
	/clone_idb="Sugano mouse liver mla"	
	/note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo (drr) primer [AATGGCCCTTTTCTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [rTGTCGACTTACTGC], digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCGCTCTTAAGCTGCG and 3' end primer CGACTGTCAGCTCGACACA."	
ORIGIN		
Query Match	27.0%; Score 86.8; DB 1;	Length 401;
Best Local Similarity	92.9%; Pred. No. 4.2e-14;	
Matches	91; Conservative 0; Mismatches=7; Indels 0; Gaps 0	
OY	219 TGGTCCACTGCCGCCGAGACACAAACCAGGAGATTGAACACTGCACACGGCCATCT	278
Db	1 TGGTCCACTGCCGCCGAGACACACCCACGAGACTTGAACACTGCACACGGCCGCTT	60
OY	279 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA	316
Db	61 GCCCAGAGAGCTGTGACACCACTTCGCTACTATCTA	98
RESULT 6	AI196314	480 bp mRNA linear EST 14-OCT-1998
LOCUS	u178h07.y1	Sugano mouse liver mla Mus musculus cDNA clone
DEFINITION	IMAGE:1887828.5; similar to gb:MS9316.mal INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579.M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.	

ACCESSION A1196314
 VERSION A1196314.1 GI:3748920
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972152
 Seq primer: custom primer used
 High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1897828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-Fl3; Site_1: DraIII (CACTGCTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCCTTTTCTTTTCTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTCTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGCTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACTCGACTCGACGACA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 4.4e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCCCCCGGAGACACAAACCCAGGACATTGAACATCTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCCCCCGGAGACACACACCCAGGACATTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACACACCACTTCGGCTACTAGCTA 316
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGGCTACTACTTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION U189F09.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:189609 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146
 VERSION A1530146.1 GI:4444281
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 Seq primer: custom primer used
 High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:189609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-Fl3; Site_1: DraIII (CACTGCTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGGCCCTTTTCTTTTCTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTCTG], digested and cloned into distinct DraIII sites of the pME18S-Fl3 vector (5' site CACTGCTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACTCGACTCGACGACA."
 ORIGIN
 Query Match 27.0%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 4.7e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 219 TGTTCACCTGCCCCCGGAGACACAAACCCAGGACATTGAACACTCTGCACACGGCCATCT 278
 DB 1 TGTTCACCTGCCCCCGGAGACACACACCCAGGACATTGAACACTGCACACGGCCGCT 60
 QY 279 GCCCAGAGAGCTGTGACACACCACTTCGGCTACTAGCTA 316
 DB 61 GCCCAGAGAGCTGTGACACCACTTCGGCTACTACTTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION U61E07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:148198 5' similar to gb:X81579 M. musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 395.
Location/Qualifiers
1..765
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1481988"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACGCTCGAGCACA."
ORIGIN
Query Match 27.0%; Score 86.8; DB 1; Length 765;
Best Local Similarity 92.9%; Pred. No. 4.8e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 219 TGTTCACCTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGACACAGGCCACTT 278
DB 1 TGTTCACCTGCCCGCGAGACACACACACCCAGGAGCATTTGAACACTGACACAGGCCACTT 60
QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 316
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 98
RESULT 9
A1530313 795 bp mRNA linear EST 18-MAR-1999
LOCUS A1530313
DEFINITION u191f01.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1889785 5' similar to gb:M59316 final INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
VERSION A1530313.1 GI:4444448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 785)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974109
Seq primer: custom primer used
High quality sequence stop: 459.
Location/Qualifiers
1..785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCACTG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGACGCTCGAGCACA."
ORIGIN
Query Match 27.0%; Score 86.8; DB 1; Length 785;
Best Local Similarity 92.9%; Pred. No. 4.8e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 219 TGTTCACCTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGACACAGGCCACTT 278
DB 1 TGTTCACCTGCCCGCGAGACACACACACCCAGGAGCATTTGAACACTGACACAGGCCACTT 60
QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 316
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCGGTACTACTA 98
RESULT 10
A1529939 799 bp mRNA linear EST 18-MAR-1999
LOCUS A1529939
DEFINITION u187c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone IMAGE:1889392 5' similar to gb:M59316 final INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION	AI529939
VERSION	AI529939.1
KEYWORDS	GI:4444074
SOURCE	EST.
ORGANISM	Mus musculus (house mouse)
REFERENCE	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 799)
AUTHORS	Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Smaller,T., Gibbons,M., Pape,D., Harvey,N., Schuck,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterson,R. and Wilson,R.
TITLE	The Mashu-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra M/Mashu-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:973716 Seq primer: custom primer used High quality sequence stop: 566. Location/Qualifiers 1..799
FEATURES	
source	

	ORIGIN	
Query Match	27.0%; Score 86.8; DB 1;	
Best Local Similarity	92.9%; Pred.No. 4.8e-14;	
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;		
QY 219 TGGTTCACCTGCCCGCCGAGACACAACCAGCAGCATTTGAACACTGTGCACACGCGCATCT 278		
Dd 1 TGATTCACATTGCCCGCGGAGACACACACCAGGAACTTGAACTGTGCACACGCGCGTCT 60		
QY 279 GCCCAGAGAGCTGTGACCAACCACTTCGCGTAATACTA 316		
Dd 61 GCCCAGAGAGCTGTGACCAACCACTTCGCGTAATACTA 98		

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Mus musculus (house mouse)	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	1 (bases 1 to 811)	NIH-MGC http://mgc.nci.nih.gov/ .	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
Contact: Robert Strausberg, Ph.D.	Email: gsabbs-remail.nih.gov	Tissue Procurement: Bradfield Laboratory	cDNA Library Preparation: Mark Bittinger	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)	DNA Sequencing by: Agencourt Bioscience Corporation	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Plate: LXM00138	row: m	column: 23	High quality sequence stop: 536.			
Location/Qualifiers						
1. .811						

ORIGIN	Query Match	27.0%; Score 86.8; DB 6; Length 811;
	Best Local Similarity	92.9%; Pred. No. 4,96-14;
	Matches	91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy	219 TGTGTCACCTGCGCCGCGAGACACAACCCGACGAGCATTTGAACACTGCACAGGCCCATCT	278
Db	39 TGTGTCACCTGCGCCGCGAGACACAACCCGACGAGCATTTGAACACTGCACAGGCCCATCT	98
Qy	279 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA	316
Db	99 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA	136
RESULT 12		
LOCUS	AI1790802	
DEFINITION	AI1790802 848 bp mRNA linear EST 02-JUL-1999	
KEYWORDS	uk28b10.y1 Sugano mouse kidney mkoa Mus musculus cDNA clone	
ACCESSION	IMAGE:1970229 5' similar to gb:X81579 M.musculus mRNA for	
VERSION	insulin-like growth factor binding (MOUSE); mRNA sequence.	
KEYWORDS	AI1790802	
SOURCE	AI1790802.1 GI:5338518	
ORGANISM	EST.	
	Mus musculus (house mouse)	

COMMENT Other ESTs: uk28b10.x1
Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039
Seq primer: custom primer used
High quality sequence stop: 514.
Location/Qualifiers
1..848
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse kidney mkl3a"
/note="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGCTG); Site 2: DraIII (CACTGCTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGCTG, 3' site CACTGCTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGG and 3' end
primer CGACTCGACGCTCGAGACA."

ORIGIN
Query Match 27.0%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 4.9e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 219 TGTTCACATGCCCCCGGAGACAAACCAGGAGATTGAACACTGCACAGGCGCATCT 278
|||||
DB 1 TGTTCACATGCCCCCGGAGACAAACCAGGAGATTGAACACTGCACAGGCGCATCT 60
QY 279 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 316
|||||
DB 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTACTACTTA 98

RESULT 13
AI528304 853 bp mRNA linear EST 18-MAR-1999
LOCUS u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone
DEFINITION IMAGE:1890210.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.
ACCESSION AI528304
VERSION AI528304.1 GI:4442439
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 853)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,T.,
Perizon,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson R.
TITLE The Mashu-NCI Mouse EST Project 1999

JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534
Seq primer: custom primer used
High quality sequence stop: 478.
Location/Qualifiers
1..852
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGCTG); Site 2: DraIII (CACTGCTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTACTGG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGCTG, 3' site CACTGCTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTCTGCTCTTAAAGCTGG and 3' end
primer CGACTCGACGCTCGAGACA."

ORIGIN
Query Match 27.0%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 4.9e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 219 TGTTCACATGCCCCCGGAGACAAACCAGGAGATTGAACACTGCACAGGCGCATCT 278
|||||
DB 1 TGTTCACATGCCCCCGGAGACAAACCAGGAGATTGAACACTGCACAGGCGCATCT 60
QY 279 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 316
|||||
DB 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTACTACTTA 98

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998
LOCUS u195d08.y1 Sugano mouse liver mla Mus musculus cDNA clone
DEFINITION IMAGE:1887663.5, similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.
ACCESSION A1196154
VERSION A1196154.1 GI:3748760
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 605)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The Mashu-NCI Mouse EST Project

JOURNAL COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:971987

FEATURES

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers
1..605

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:187663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTT] (TTTTTTT); double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTTCTCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 26.5%; Score 85.2; DB 1; Length 605;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 219 TGGTCCACTGCCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACAGGCGCATCT 278
|||||
DB 1 TGGTCCACTGCCCGCGAGACACACACCCAGAGCATTTGAACACTGCACAGGCGCATCT 60
|||||
QY 279 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||||
DB 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98
|||||
```

RESULT 15

AT785039

LOCUS 380 bp mRNA linear EST 02-JUL-1999
DEFINITION ui73a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1888018.5' similar to gb:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AT785039
VERSION AT785039.1 GI:5332755

KEYWORDS

EST.
Mus musculus
Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 380)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.

AUTHORS

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)

COMMENT

Other ESTs: ui73a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:972342

FEATURES

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers
1..380

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTT] (TTTTTTT); double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTTCTCTCTAAAGCTGCG and 3' end
primer CGACTCGAGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 25.6%; Score 82.2; DB 1; Length 380;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```
QY 222 TCCACTGCCCGCGAGACACAAACCCAGAGCATTTGAACACTGCACAGGCGCATTTGCC 281
|||||
DB 1 TCCACTGCCCGCGAGAGACACACACCAGAGCATTTGAACACTGCATGCGCCGCTGCC 60
|||||
QY 282 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 316
|||||
DB 61 CAGAGAGCTGTGACCACTTCGCTACTAGCTA 95
|||||
```

Search completed: September 1, 2005, 06:56:49
Job time : 241.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 934.893 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916B-4

Perfect score: 321

Sequence: 1 tacactggggggcagagtcgc.....ttccgctactagctagccgc 321

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	321	100.0	321	9	US-09-972-916A-4
2	321	100.0	372	9	US-09-972-916A-5
3	219.4	66.3	270	9	US-09-972-916A-3
4	219.4	66.3	423	9	US-09-972-916A-6
5	219	66.2	219	9	US-09-972-916A-2
6	103	33.1	423	9	US-09-972-916A-6
7	76.4	23.8	1500	9	US-09-917-800A-1608

8	76.4	23.8	1500	17	US-10-388-934-5	Sequence 5, Appli
9	76.4	23.8	1500	17	US-10-191-803-73	Sequence 73, Appl
10	76.4	23.8	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
11	63.8	19.9	6128	9	US-09-880-107-2393	Sequence 2393, Ap
12	63.8	19.9	6128	22	US-10-756-149-1484	Sequence 1484, Ap
13	63.8	19.9	9173	22	US-10-893-315-126	Sequence 126, App
14	63.8	19.9	9174	22	US-10-893-315-160	Sequence 160, App
15	55.2	17.2	13011	17	US-10-388-934-36	Sequence 36, Appli
16	55.2	17.2	13011	18	US-10-152-319A-2130	Sequence 2150, Ap
17	51	15.9	51	9	US-09-972-916A-1	Sequence 1, Appli
18	51	15.9	270	9	US-09-972-916A-3	Sequence 3, Appli
19	40.2	12.5	31703	17	US-10-085-117-172	Sequence 172, App
20	37.8	11.8	7061	15	US-10-311-655-970	Sequence 970, App
21	37.8	11.8	7061	17	US-10-221-613-148	Sequence 148, App
22	36.6	11.4	1678	17	US-10-094-749-9	Sequence 9, Appli
23	36	11.2	761	17	US-10-264-237-828	Sequence 828, App
24	35.6	11.1	1487	20	US-09-918-995-22070	Sequence 22070, A
25	35.2	11.0	1687	20	US-10-739-930-5461	Sequence 5461, Ap
26	35	10.9	769	13	US-10-027-632-164336	Sequence 164336,
27	35	10.9	769	17	US-10-027-632-164336	Sequence 164336,
28	35	10.9	3895	14	US-10-011-585A-76	Sequence 76, Appli
29	34	10.6	49979	19	US-10-741-601-5746	Sequence 5746, Ap
30	34	10.6	49979	21	US-10-741-601-17905	Sequence 17905, A
31	32.8	10.2	2130	9	US-09-909-962A-7	Sequence 7, Appli
32	32.8	10.2	2130	9	US-09-909-962A-7	Sequence 7, Appli
33	32.4	10.1	1389	19	US-10-437-863-79182	Sequence 79182, A
34	32.4	10.1	1584	19	US-10-437-863-79185	Sequence 79185, A
35	32.4	10.1	30000	11	US-09-980-217-1	Sequence 1, Appli
36	32.2	10.0	750	13	US-10-027-632-25562	Sequence 25562, A
37	32.2	10.0	750	13	US-10-027-632-25562	Sequence 25562, A
38	32.2	10.0	750	13	US-10-027-632-25562	Sequence 25562, A
39	32.2	10.0	750	17	US-10-027-632-25562	Sequence 25562, A
40	32.2	10.0	750	17	US-10-027-632-25563	Sequence 25563, A
41	32.2	10.0	750	17	US-10-027-632-25563	Sequence 25563, A
42	32.2	10.0	755	13	US-10-027-632-150407	Sequence 150407,
43	32.2	10.0	755	17	US-10-027-632-150407	Sequence 150407,
44	32	10.0	76180	19	US-10-322-281-492	Sequence 492, App
45	32	10.0	91760	13	US-10-087-192-844	Sequence 844, App

ALIGNMENTS

RESULT 1
US-09-972-916A-4
Sequence 4, Application US/09972916A
Patent No. US20020107198A1
GENERAL INFORMATION:
APPLICANT: Thule, Peter M.
TITLE OF INVENTION: GUCCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
FILE REFERENCE: US 1292/01 (VA)
CURRENT APPLICATION NUMBER: US/09/972, 916A
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/239, 113
PRIOR FILING DATE: 2000-10-11
NUMBER OF SEQ ID NOS: 6
SEQ ID NO 4
LENGTH: 321
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 100.0%; Score 321; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 9, 2e-100; Indels 0; Gaps 0;
Matches 321; Conservative 0; Mismatches 0;

QY 1 TACACTGGGGGCCAGAGTCGAGAACCAACGAGAGTCCCCGTGCGCCCATGTACACTGGG 60
DB 1 TACACTGGGGGCCAGAGTCGAGAACCAACGAGAGTCCCCGTGCGCCCATGTACACTGGG 60
QY 61 GGCGAGTGCAGGAGACCAACGAGAGTCCCCGTGCGCCCATGTACAGCAAGAAACAAACT 120

```
Db      61 GGCACAGATCCAGGAACCAAGGAGTGGCCCGTGCCTCATGTCAACAAACAACT 120
        |||
Qy      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCATTAACCCGTCTGCCAG 180
        |||
Db      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCATTAACCCGTCTGCCAG 180
        |||
Qy      181 CCAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATGTCTCACTGCCCGCAGACA 240
        |||
Db      181 CCAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATGTCTCACTGCCCGCAGACA 240
        |||
Qy      241 CAAACCCAGCGAGCATTTGAACACTGACACGCGCATCTGCCCAAGAGCTGTGACCA 300
        |||
Db      241 CAAACCCAGCGAGCATTTGAACACTGACACGCGCATCTGCCCAAGAGCTGTGACCA 300
        |||
Qy      301 CTTCCGCTACTAGCTAGCCGC 321
        |||
Db      301 CTTCCGCTACTAGCTAGCCGC 321
        |||
```

RESULT 2
US-09-972-916A-5

```
/ Sequence 5, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ CURRENT FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 5
/ LENGTH: 372
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-5
```

```
Query Match      100.0%; Score 321; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 9.5e-100;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 TACACTGGGGGCGCAGAGTCCAGGAACCAAGGAGTGGCCCGTGCCTCATGTACACTGG 60
        |||
Db      52 TACACTGGGGGCGCAGAGTCCAGGAACCAAGGAGTGGCCCGTGCCTCATGTACACTGG 111
        |||
Qy      61 GGCAGAGTCCAGGAACCAAGGAGTGGCCCGTGCCTCATGTCAACAAACAAACT 120
        |||
Db      112 GGCAGAGTCCAGGAACCAAGGAGTGGCCCGTGCCTCATGTCAACAAACAAACT 171
        |||
Qy      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCATTAACCCGTCTGCCAG 180
        |||
Db      121 TATTTGAACAGGGGATCTAGACAGCTGCCCTGACATCATTAACCCGTCTGCCAG 231
        |||
Qy      181 CCAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATGTCTCACTGCCCGCAGACA 240
        |||
Db      232 CCAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATGTCTCACTGCCCGCAGACA 291
        |||
Qy      241 CAAACCCAGCGAGCATTTGAACACTGACACGCGCATCTGCCCAAGAGCTGTGACCA 300
        |||
Db      241 CAAACCCAGCGAGCATTTGAACACTGACACGCGCATCTGCCCAAGAGCTGTGACCA 351
        |||
Qy      301 CTTCCGCTACTAGCTAGCCGC 321
        |||
Db      352 CTTCCGCTACTAGCTAGCCGC 372
        |||
```

RESULT 3

```
US-09-972-916A-3
/ Sequence 3, Application US/09972916A
```

```
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ CURRENT FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 3
/ LENGTH: 270
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-3
```

```
Query Match      68.3%; Score 219.4; DB 9; Length 270;
Best Local Similarity 99.5%; Pred. No. 6.7e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      101 TGTCAACGCAAAACAACTTATTTGAACAGGGGATCTTAGACAGCTGCCCTGACAA 160
        |||
Db      50 TATCAACGCAAAACAACTTATTTGAACAGGGGATCTTAGACAGCTGCCCTGACAA 109
        |||
Qy      161 CATTAACCCGCTGCGGAGCGAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATG 220
        |||
Db      110 CATTAACCCGCTGCGGAGCGAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATG 169
        |||
Qy      221 GTCCACTGCGCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACAGCCCATCTGC 280
        |||
Db      170 GTCCACTGCGCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACAGCCCATCTGC 229
        |||
Qy      281 CCAGAGAGCTGTGACACCACTTCGCTACTAGCTAGCCGC 321
        |||
Db      230 CCAGAGAGCTGTGACACCACTTCGCTACTAGCTAGCCGC 270
        |||
```

RESULT 4

```
US-09-972-916A-6
/ Sequence 6, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ CURRENT FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 6
/ LENGTH: 423
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-6
```

```
Query Match      68.3%; Score 219.4; DB 9; Length 423;
Best Local Similarity 99.5%; Pred. No. 7.5e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      101 TGTCAACGCAAAACAACTTATTTGAACAGGGGATCTTAGACAGCTGCCCTGACAA 160
        |||
Db      203 TATCAACGCAAAACAACTTATTTGAACAGGGGATCTTAGACAGCTGCCCTGACAA 262
        |||
Qy      161 CATTAACCCGCTGCGGAGCGAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATG 220
        |||
Db      263 CATTAACCCGCTGCGGAGCGAGCCCTTCATTAAGGCGCTGGGTATGCGCAGCCAGCATG 322
        |||
Qy      221 GTCCACTGCGCGCGAGACACAAACCCAGCAGCATTTGAACACTGACACAGCCCATCTGC 280
        |||
```

Db 323 GTCACATGCCCCGAGACACAAACCGAGCATTTGAACACTGACACGCGCATCTCC 382
Qy 281 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTACGCCG 321
Db 383 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTACGCCG 423

RESULT 5

US-09-972-916A-2
/ Sequence 2, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 2
/ LENGTH: 219
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
US-09-972-916A-2

Query Match 68.2%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 8.7e-65;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 TCACAAGCAAAACAATTATTTTGAACACGGGATCTTGAACACGCTGCTGACATCA 162
Db 1 TCACAAGCAAAACAATTATTTTGAACACGGGATCTTGAACACGCTGCTGACATCA 60
Qy 163 TTACCCGCTGCGAGCGAGCCCTTCATAGGCCCTGGTATGCGCCAGCCAGATGCT 222
Db 61 TTACCCGCTGCGAGCGAGCCCTTCATAGGCCCTGGTATGCGCCAGCCAGATGCT 120
Qy 223 CCAGTCCCGCCGAGACACAAACCGAGCATTTGAACATGACACAGCGCATCTGCC 282
Db 121 CCAGTCCCGCCGAGACACAAACCGAGCATTTGAACATGACACAGCGCATCTGCC 180
Qy 283 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACGCCG 321
Db 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACGCCG 219

RESULT 6

US-09-972-916A-6/c
/ Sequence 6, Application US/09972916A
/ Patent No. US20020107198A1
/ GENERAL INFORMATION:
/ APPLICANT: Thule, Peter M.
/ TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
/ FILE REFERENCE: US 1292/01 (VA)
/ CURRENT APPLICATION NUMBER: US/09/972,916A
/ PRIOR FILING DATE: 2001-10-10
/ PRIOR APPLICATION NUMBER: US 60/239,113
/ PRIOR FILING DATE: 2000-10-11
/ NUMBER OF SEQ ID NOS: 6
/ SEQ ID NO 6
/ LENGTH: 423
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthesized
US-09-972-916A-6

Query Match 32.1%; Score 103; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 6.9e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCACTGGGGGCGACAGTTCAGAAACAGGGAGTCCCCGTCGCCCATTTACACTGGG 60
Db 204 TCACTGGGGGCGACAGTTCAGAAACAGGGAGTCCCCGTCGCCCATTTACACTGGG 145
Qy 61 GGCAGAGTCCAGAAACAGGGAGTGGCCCGTGGCCCATGT 103
Db 144 GGCAGAGTCCAGAAACAGGGAGTGGCCCGTGGCCCATGT 102

RESULT 7

US-09-917-800A-1608
/ Sequence 1608, Application US/09917800A
/ Patent No. US20020119462A1
/ GENERAL INFORMATION:
/ APPLICANT: Mendrick, Donna
/ APPLICANT: Porter, Mark
/ APPLICANT: Johnson, Kory
/ APPLICANT: Castle, Arthur
/ APPLICANT: Elashoff, Michael
/ TITLE OF INVENTION: Molecular Toxicology Modeling
/ FILE REFERENCE: 44921-5038-US
/ CURRENT APPLICATION NUMBER: US/09/917,800A
/ PRIOR FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,040
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: US 60/222,880
/ PRIOR FILING DATE: 2000-11-02
/ PRIOR APPLICATION NUMBER: US 60/290,029
/ PRIOR FILING DATE: 2001-05-11
/ PRIOR APPLICATION NUMBER: US 60/290,645
/ PRIOR FILING DATE: 2001-05-15
/ PRIOR APPLICATION NUMBER: US 60/292,336
/ PRIOR FILING DATE: 2001-05-22
/ PRIOR APPLICATION NUMBER: US 60/295,798
/ PRIOR FILING DATE: 2001-06-06
/ PRIOR APPLICATION NUMBER: US 60/297,457
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,884
/ PRIOR FILING DATE: 2001-06-19
/ PRIOR APPLICATION NUMBER: US 60/303,459
/ PRIOR FILING DATE: 2001-07-09
/ NUMBER OF SEQ ID NOS: 1740
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 1608
/ LENGTH: 1500
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
/ FEATURE:
/ OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144
US-09-917-800A-1608

Query Match 23.8%; Score 76.4; DB 9; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 239 CACAAACCCAGCGAGCATTTGAACACTGACACAGCGCATCTGCCAGAGAGCTGTGACCAC 298
Db 8 CACAAACCCAGCGAGCATTTGAACACTGACACAGCGCATCTGCCAGAGAGCTGTGACCAC 67
Qy 299 CACTTCGCTACTAGCTA 316
Db 68 CACTTCGCTACTAGCTA 85

RESULT 8

US-10-386-934-5
/ Sequence 5, Application US/10386934
/ Publication No. US20040005547A1
/ GENERAL INFORMATION:
/ APPLICANT: Boess, Franziska
/ APPLICANT: Suter-Dick, Laura

```

; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US2004000547Alway rat)
US-10-388-934-5

Query Match      23.8%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 298
      |||
Db      8 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 67

Qy      239 CACTTCGGCTACTACTA 316
      |||
Db      68 CACTTCGGCTACTACTA 85

RESULT 9
US-10-191-803-73
; Sequence 73, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_013144
US-10-191-803-73

Query Match      23.8%; Score 76.4; DB 17; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 298
      |||
Db      8 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 67

Qy      239 CACTTCGGCTACTACTA 316
      |||
Db      239 CACTTCGGCTACTACTA 316
```

```

Db      68 CACTTCGGCTACTACTA 85

RESULT 10
US-10-152-319A-1613
; Sequence 1613, Application US/10152319A
; Publication No. US20040072160A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Higgs, Brandon
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5089-US
; CURRENT APPLICATION NUMBER: US/10/152,319A
; CURRENT FILING DATE: 2002-05-22
; PRIOR APPLICATION NUMBER: US 60/292,335
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/297,523
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,925
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,810
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,807
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/303,808
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/315,047
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/324,928
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/330,867
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: US 60/330,462
; PRIOR FILING DATE: 2001-10-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2221
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1613
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match      23.8%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.3e-15;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      239 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 298
      |||
Db      8 CACAACCAGGAGCATTTGAACACTGACACGCGCATCTGCCAGAGAGCTGTGACCAC 67

Qy      239 CACTTCGGCTACTACTA 316
      |||
Db      68 CACTTCGGCTACTACTA 85

RESULT 11
US-09-880-107-2393
; Sequence 2393, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
```



```

; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2393
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
US-09-880-107-2393

Query Match          19.9%; Score 63.8; DB 9; Length 6128;
Best Local Similarity 63.2%; Pred. No. 3.8e-11;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

69  TCCAGGAACCAAGGAGTCCCGGCGCCATGTCAACAAGCAAACTTATTTTGA 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438  TCTTCCCAACGAGCGGTTTGGTGAAGGCGCTTGAGTCAAGAAACAACTTATTTTGA 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129  ACACGGGAGTCTAGACGCGTGCCTGACATCATTAACCC-----GTGCTGCCG 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498  AACATCAGCTCTTAGCGTGGCGGCTGCGCAATCATTAACCTCTGTCAGAGTGGCGCG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
179  AGCCAGCCCTTATAGGCGCTGATGAGCCAGCAGCAGATGTCACATGCGCGCGAGA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
558  CCGTGGCCCTTTATAGGTCGGCGGTGTCCAGAGAGATGGCGACCGCATCC---- 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239  CACAAACCAAGGAGATGAACACTGC-ACACGCGCATCTGCGCCAGAGAGCTGTGACCA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
614  ----CATCCAGGAGCATCTGCGCGCGCGCGCGCCGACCTCCAGAGAGACTGGCCA 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298  CCACTTC 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
670  CCGCTCC 676

RESULT 12
US-10-756-149-1484
; Sequence 1484, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natsasha
; APPLICANT: Zlocnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1484
; LENGTH: 6128
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-756-149-1484

Query Match          19.9%; Score 63.8; DB 22; Length 6128;
Best Local Similarity 63.2%; Pred. No. 3.8e-11;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

69  TCCAGGAACCAAGGAGTCCCGGCGCCATGTCAACAAGCAAACTTATTTTGA 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
438  TCTTCCCAACGAGCGGTTTGGTGAAGGCGCTTGAGTCAAGAAACAACTTATTTTGA 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129  ACACGGGAGTCTAGACGCGTGCCTGACATCATTAACCC-----GTGCTGCCG 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498  AACATCAGCTCTTAGCGTGGCGGCTGCGCAATCATTAACCTCTGTCAGAGTGGCGCG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
129  ACACGGGAGTCTAGACGCGTGCCTGACATCATTAACCC-----GTGCTGCCG 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498  AACATCAGCTCTTAGCGTGGCGGCTGCGCAATCATTAACCTCTGTCAGAGTGGCGCG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 179 AGCCAGCCCTTCATAGGCGCTGGTATGCGCAGCAGATGTCACATGCGCGCGAGA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 558 CCGTGGCCCTTTATAGGTCGGCGGTGTCCAGAGCATGCGCACCGCATCC---- 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CACAAACCAAGGAGATGAACACTGC-ACACGCGCATCTGCCAGAGAGCTGTGACCA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 614 ----CATCCAGGAGCATCTGCGCGCGCGCGCGCGCCACCTCCAGAGAGACTGGCCA 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 CCACTTC 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 670 CCGCTCC 676

RESULT 13
US-10-893-315-126
; Sequence 126, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; PRIOR FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
; NUMBER OF SEQ ID NOS: 2172
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-10-893-315-126

Query Match          19.9%; Score 63.8; DB 22; Length 9173;
Best Local Similarity 63.2%; Pred. No. 4.3e-11;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

69  TCCAGGAACCAAGGAGTCCCGGCGCCATGTCAACAAGCAAACTTATTTTGA 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1842 TCTTCCCAACGAGCGGTTTGGTGAAGGCGCTTGAGTCAAGAAACAACTTATTTTGA 1901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 ACACGGGAGTCTAGACGCGTGCCTGACATCATTAACCC-----GTGCTGCCG 178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1902 AACATCAGCTCTTAGCGTGGCGGCTGCGCAATCATTAACCTCTGTCAGAGTGGCGCG 1961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 179 AGCCAGCCCTTATAGGCGCTGATGAGCCAGCAGCAGATGTCACATGCGCGCGAGA 238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1962 CCGTGGCCCTTTATAGGTCGGCGGTGTCCAGAGCATGCGCACCGCATCC---- 2017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 239 CACAAACCAAGGAGATGAACACTGC-ACACGCGCATCTGCCAGAGAGCTGTGACCA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2018 ----CATCCAGGAGCATCTGCGCGCGCGCGCGCGCCACCTCCAGAGAGACTGGCCA 2073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 CCACTTC 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2074 CCGCTCC 2080

RESULT 14
US-10-893-315-160
; Sequence 160, Application US/10893315
; Publication No. US20050147987A1
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
; FILE REFERENCE: CL000786
; CURRENT APPLICATION NUMBER: US/10/893,315
; CURRENT FILING DATE: 2004-07-19
; PRIOR APPLICATION NUMBER: 60/231,397
```

```

/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 2172
/ SOFTWARE: PatSeq for Windows Version 4.0
/ SEQ ID NO: 160
/ LENGTH: 9174
/ TYPE: DNA
/ ORGANISM: Human
US-10-893-315-160

```

Query Match	19.9%	Score 63.8	DB 22	Length 9174
Best Local Similarity	63.2%	Pred.No.4.3e-11		
Matches 156; Conservative	0	Mismatches 72	Indels 19	Gaps 3

QY 69 TCACGGAACCCGGAGTGCCTCCGCGCCCATGTACAAAGCAAAACAAATTATTTTGA 128
Db 1842 TCTCTCCACACGCGCTTTTCGTAGGGCTTGGGTGCATCAGCAAAACAAACTTATTTTGA 1900

QY 129 ACACGGGGATCCTAGCACGCTGCCCTGAACATCATTTAACCC-----GTGCTGCC 178
Db 1902 ACACATCAGCTCCACGCTGCGCGCTGCGCAATTCATTAACCTCCTGTGCAAGTGGCGCG 1961

0y AGCAGGCCCTTCAATAGGCCCTGGGTAATGCCAGCAGCATGTCACCTGCCCGCGAGA 238
Db 1962 CCTGTGCCCTTATATAGGCGCGCGCTGTATCCAGCAGAGATGGCGCACCGGCATTC--- 2017

QY 239 CACAAACCCAGCGACATTGAACACTGC-AACCGGCATCTGCCAGAGACTGTGACCA 297
Db 2018 ----CATCCAGCGACATCTGCCCGCGCGCGCCACACTCTCCAGAGAGCACTGTGCCA 2073

Qy	298	CCACTTC	304
Db	2074	CCGCTCC	2080

RESULT 15
US-10-388-934-36/c

```

; Publication No. US20040005547A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Boess, Franziska

```

```

; APPLICANT: wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199

```

CURRENT FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: 02005336.9
PRIOR FILING DATE: 2002-03-14

```

;
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ. ID NOS: 862
;
; SOFTWARE: PatentIn version 3.1

```

```

;
; LENGTH: 13011
;
; TYPE: DNA
;
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)

```

Query Match	17.2%	Score 55.2;	DB 17;	Length 13011;
Best Local Similarity	88.2%	Pred. No. 4.2e-08;		

35 TGGCCCGGTGGCCCATGTACCTGGGGGCCAGAGTCCAGAAACACGAGGTGCCCCG 94

[illegible]

Search completed: September 1, 2005, 16:44:28

; Patent No. 6825336

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMOPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL000788
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match 19.9%; Score 63.8; DB 4; Length 9174;
Best Local Similarity 63.2%; Pred. No. 1.3e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 69 TCCAGAACCCAGGAGTGCCCGCTGCGCCCATGTCCAGACAGAAACAACTTATTGTA 128
DB 1842 TCCCTCCACACGAGGTTGCGTAGGCGCTTGCGTGCATGACAAACAACTTATTGTA 1901
QY 129 ACACGGGAGTCTAGCAGCGCTGCCGTGACATCATTAACCC-----GTGCTGCCG 178
DB 1902 ACACCTAGCTCTAGGCTGCGCGCGCTGCAATCATTAACCTCTGTGCAAGTGGCGG 1961
QY 179 AGCCAGCCCTTATTAAGCCCTGGGTANGGCGACGAGCATGTCACCTGCCCGCGAGA 238
DB 1962 CCGTGCCTTATTAAGTGCCTGCTGTGTCAGAGCATGCGCACCGCATCC---- 2017
QY 239 CACAACCCAGGAGATGAAACATGTC-ACAAGGCATCTGCCAGAGAGCTGTGACCA 297
DB 2018 ----CATTCAGGAGATCTGCCGCGCGCGCCGCCACCTCCAGAGAGCATGGCCA 2073
QY 298 CCACTTC 304
DB 2074 CCGCTCC 2080

RESULT 3
US-08-945-140-1/c
Sequence 1, Application US/08945140
Patent No. 6109878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSES: Rhone-Poulenc Rorer Inc.,
STREET: 500 Arcola Road, Mailstop 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Esq., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match 17.2%; Score 55.2; DB 3; Length 194;
Best Local Similarity 88.2%; Pred. No. 1.6e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCGTGCGCCCATGTGACACTGGGGCCAGAGTCCAGAAACCGGAGTCCCGCTG 94
DB 78 TGCCAACGGAAGCCTTGTAACACTGGGGGCGAGTCCAGAAACCGGAGTCCCGCTG 19
QY 95 CGCCCATG 102
DB 18 CGCCCATG 11

RESULT 4
US-08-791-849A-14/c
Sequence 14, Application US/08791849A
Patent No. 591449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
TITLE OF INVENTION: Lipid Content in Plant Seed
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSES: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
FEATURE:
NAME/KEY: intron
LOCATION: join(3219..3765, 3949..5916, 6009..6151,
LOCATION: 6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
LOCATION: 9480..10162)
US-08-791-849A-14
```

```
Query Match 17.2% Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.1e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 35 TGCCCGCGTGCAGTGTACCTGAGTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 94
DB 3088 TCCCAACGAAAGCTTGTACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 3029
QY 95 CGCCCATG 102
DB 3028 CGCCCATG 3021
```

```
RESULT 5
US-09-949-016-11786/c
Sequence 11786, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11786
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-11786
```

```
Query Match 12.5% Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.011;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 34 GTGCCCCGTGCGCCATGTACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 93
DB 2417 GTTCCATTGGAAGCCCTGTATGCGAGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357
```

```
RESULT 6
US-09-949-016-17205/c
```

```
Sequence 17205, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17205
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-17205
```

```
Query Match 12.5% Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.011;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

```
QY 34 GTGCCCCGTGCGCCATGTACACTGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 93
DB 2417 GTTCCATTGGAAGCCCTGTATGCGAGGGCCAGAGTCCAGAAACACGGAGTGCCTCGT 2358
QY 94 G 94
DB 2357 G 2357
```

```
RESULT 7
US-09-484-970B-138/c
Sequence 138, Application US/09484970B
Patent No. 6426186
GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmueth, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 138
LENGTH: 3054
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. 6426186 336987.1CB1
US-09-484-970B-138
```

```
Query Match 11.4% Score 36.6; DB 3; Length 3054;
Best Local Similarity 62.6%; Pred. No. 0.087;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 2 AACTGGGGGCCAGAGTCCAGAAACACGGAGTGCCTCGCCCATGTACTGAGG 61
DB 355 ACTTCGGAAGCAAGACCCCTGTATGCGAGGGGTAGCTCGAGCGTCCATCCAGCCGGTG 296
QY 62 GCCAGATCCAGAAACACGGAGTGCCTCGG 92
DB 295 GCTGAGGCTTGGAGCGAAGGAGAGCCCGG 265
```

```
RESULT 8
```

US-09-949-016-19194/c
; Sequence 19194, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19194
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-19194

Query Match 10.9%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGTCACGTGCCCCGACGACA 240
DB 478 CCAGGCTTCCTCGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGACAGTTAGGCC 419
QY 241 CAACCCGAGCAGCATTTGAACATGTCACAGCGGCATCTGCGAGAGGTGACCA 300
DB 418 CCCATCCAGAGACACATCCACCCGATCCAGAGCCTTGCACCAACCATTCATGATGACC 359
QY 301 CTT 303
DB 358 CTT 356

RESULT 9
US-09-949-016-73784/c
; Sequence 73784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73784
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-73784

Query Match 10.9%; Score 35; DB 4; Length 601;
Best Local Similarity 55.3%; Pred. No. 0.14;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGTCACGTGCCCCGACGACA 240
DB 478 CCAGGCTTCCTCGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGACAGTTAGGCC 419

QY 241 CAACCCGAGCAGCATTTGAACATGTCACAGCGGCATCTGCGAGAGGTGACCA 300
DB 418 CCCATCCAGAGACACATCCACCCGATCCAGAGCCTTGCACCAACCATTCATGATGACC 359
QY 301 CTT 303
DB 358 CTT 356

RESULT 10
US-09-949-016-2153/c
; Sequence 2153, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2153
; LENGTH: 2393
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2153

Query Match 10.9%; Score 35; DB 4; Length 2393;
Best Local Similarity 55.3%; Pred. No. 0.26;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAGCCCTTCATTAAGCCCTGGGTATGCGCCAGCCAGCATGTGTCACGTGCCCCGACGACA 240
DB 2182 CCAGGCTTCCTCGGGGCGCTGTCTCTCCATCCAGCTTACGACCTGCGACAGTTAGGCC 2123
QY 241 CAACCCGAGCAGCATTTGAACATGTCACAGCGGCATCTGCGAGAGGTGACCA 300
DB 2122 CCCATCCAGAGACACATCCACCCGATCCAGAGCCTTGCACCAACCATTCATGATGACC 2063
QY 301 CTT 303
DB 2062 CTT 2060

RESULT 11
US-09-949-016-74/c
; Sequence 74, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 2408

TYPE: DNA
ORGANISM: Human
US-09-949-016-74

Query Match 10.9%; Score 35; DB 4; Length 2408;
Best Local Similarity 55.3%; Pred. No. 0.27;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 2190 CCAAGCCCTTCCTGGGGCGCTGTCTCTCTCCTCCTCAGCTTACCTGACCTGCGAGATTAGGCC 2131
QY 241 CAAACCCAGCAGCATTTGAACACTGCACGCGCATCTGCCAGAGAGCTGTGACACCA 300
DB 2130 CCCATCCAGAGACATCCACCAGATCCAGGCTTCCACACCACTCCATGATGATGCC 2071
QY 301 CTT 303
DB 2070 CTT 2068

RESULT 12

US-09-949-016-11816/c
Sequence 11816, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11816
LENGTH: 7561
TYPE: DNA
ORGANISM: Human
US-09-949-016-11816

Query Match 10.9%; Score 35; DB 4; Length 7561;
Best Local Similarity 55.3%; Pred. No. 0.45;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 5350 CCAAGCCCTTCCTGGGGCGCTGTCTCTCCTCCTCAGCTTACCTGACCTGCGAGATTAGGCC 5291
QY 241 CAAACCCAGCAGCATTTGAACACTGCACGCGCATCTGCCAGAGAGCTGTGACACCA 300
DB 5290 CCCATCCAGAGACATCCACCAGATCCAGGCTTCCACACCACTCCATGATGATGCC 5231
QY 301 CTT 303
DB 5230 CTT 5228

RESULT 13

US-09-949-016-13895/c
Sequence 13895, Application US/09949016

Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13895
LENGTH: 7562
TYPE: DNA
ORGANISM: Human
US-09-949-016-13895

Query Match 10.9%; Score 35; DB 4; Length 7562;
Best Local Similarity 55.3%; Pred. No. 0.45;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 181 CCAAGCCCTTCATTAAGCCCTGGGTATGCGCAGCCAGCATGTCCTGCCCCCGAGACA 240
DB 5350 CCAAGCCCTTCCTGGGGCGCTGTCTCTCCTCCTCAGCTTACCTGACCTGCGAGATTAGGCC 5291
QY 241 CAAACCCAGCAGCATTTGAACACTGCACGCGCATCTGCCAGAGAGCTGTGACACCA 300
DB 5290 CCCATCCAGAGACATCCACCAGATCCAGGCTTCCACACCACTCCATGATGATGCC 5231
QY 301 CTT 303
DB 5230 CTT 5228

RESULT 14

US-09-270-767-29006
Sequence 29006, Application US/09270767

Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 29006
LENGTH: 364
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-29006

Query Match 10.4%; Score 33.4; DB 4; Length 364;
Best Local Similarity 51.2%; Pred. No. 0.38;
Matches 105; Conservative 0; Mismatches 96; Indels 4; Gaps 1;

QY 82 GGAAGGCCCGGCGCCCATGTCAACAGCAAAACAACTTATTGGAACAGCGGGA----- 137
DB 160 GAGAGCTTATGCGCGCCCATGCAACAGTACCGCCATGATGTTCGACAAAGGGGACGCG 219
QY 138 TCCTAGCAGCGCTGCGCCTGACATATTAACCGGTGCTGCCAGCGCCCTTCATTAAGGC 197
DB 220 GCCGAACAGACTTGTGTGCGCATGCTACCGCCCTGTGATCATGTCCGCTATTATACGGGC 279
QY 198 CTTGGGTATGCGCAGCAGCATGATTCACCTGCCCGCGAGACACAAACCCAGCAGCATT 257
DB 280 CTTAACAGGAGAAAGAAAGAGCCAGACCACTAGCGGAAACAAACCCACCAACAT 339
QY 258 GAACACTGACACCGCCATCTGCCC 282
DB 340 GCAACAGCATTCAGCCACGATCC 364

RESULT 15

US-09-270-767-13109

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: August 31, 2005, 13:10:38 ; Search time 2316.24 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 tacactggggggccagagatcc.....ttccgctactagctagccgc 372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hhg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sbs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	214.4	57.6	1181	10	RAT1GFZ
2	214.4	57.6	185148	2	AC136382
3	188.8	50.8	1363	10	MM1LGF
4	188.8	50.8	192843	10	AL607124
5	98.2	26.4	1569	10	BC013345
6	79.2	21.3	1507	10	AY560836
7	76.4	20.5	1500	6	AX401932
8	76.4	20.5	1500	6	AX827271
9	76.4	20.5	1500	10	RAT1GFB
10	76.4	20.5	5001	6	AX163782
11	76.4	20.5	5001	10	RAT1GFB
12	65.4	17.6	1510	10	BC078889
13	64	17.2	3886	9	AY095345
14	63.8	17.2	6128	6	AX409747
15	63.8	17.2	6128	9	HUM1GFBP1A
16	63.8	17.2	6480	11	G19994
17	63.8	17.2	6480	9	HUM1GFBP1
18	63.8	17.2	9082	9	AY434089
19	63.8	17.2	69887	9	AC091524

C 20	62.2	16.7	141539	9	AC146152
C 21	62.2	16.7	189932	2	AC146117
C 22	62.2	16.7	200935	2	AC148834
C 23	60.8	16.3	448	11	G67139
C 24	55.2	14.8	194	6	A57715
C 25	55.2	14.8	194	6	AR175909
C 26	55.2	14.8	13011	6	E14395
C 27	55.2	14.8	13011	6	AX827302
C 28	55.2	14.8	13011	10	RAT1GFB
C 29	54.4	14.6	599	9	HUM1GFBP1
C 30	52	14.0	835	11	G67173
C 31	47.4	12.7	2717	2	RAT1GFB
C 32	47.4	12.7	231241	2	AC097039
C 33	44.6	12.0	184889	9	AL442125
C 34	44.6	12.0	256781	2	AC097952
C 35	44.6	12.0	259329	2	AC109570
C 36	44.2	11.9	1480	10	RAT1GFBP1
C 37	44.2	11.9	206515	10	AC140332
C 38	44.2	11.9	215533	2	AC149086
C 39	42.6	11.5	43295	2	AC006177
C 40	42.6	11.5	53370	9	AL592071
C 41	42.6	11.5	186314	10	AC003694
C 42	42.4	11.4	170654	10	AC116502
C 43	42.4	11.4	187397	2	AC114566
C 44	41.4	11.1	200340	9	AC100797
C 45	41	11.0	81704	9	AC110299

ALIGNMENTS

RESULT 1
LOCUS RAT1GFZ
DEFINITION Rattus norvegicus insulin-like growth factor gene fragment.
ACCESSION M84484
VERSION M84484.1 GI:204927
KEYWORDS insulin-like growth factor.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
1 (bases 1 to 1181)
Unterman,T.G., Jackson,R.G., McGary,E., Whalen,C. and Goswami,R.G.
Biochem. Biophys. Res. Commun. (1991) In press
COMMENT
Original source text: Rattus norvegicus (strain Sprague-Dawley)
male adult liver DNA.

FEATURES

source location/Qualifiers
1..1181
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="liver"
/dev_stage="adult"

ORIGIN

Query Match 57.6%; Score 214.4; DB 10; Length 1181;
Baser Local Similarity 97.3%; Pred. No. 1.5e-51;
Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
144 TGGCCCATGTGCACAGCAAACTATTGTGAACACGGAGATCTTACAGCGTCC 203
|||||
813 TGTGGGAGCTCACAAGCAAACTATTGTGAACACGGAGATCTTACAGCGTCC 872
|||||
204 CTGACATCATTAACCGTGTGCGGAGCCGCTTATAGGCGCTTGAGTGGCCAG 263
|||||
873 CTGACATCATTAACCGTGTGCGGAGCCGCTTATAGGCGCTTGAGTGGCCAG 932
|||||
264 CCAGCATGTCACATGCGCGCGAGACACCAACCGAGCATTTGAACACTGACACAG 323
|||||

Db 993 CCAGCATGGTCCAGCTGCCCGCCGAGACACAAACCAGCATGTAACACTGCACACGG 992
Qy 324 CCATCTCCCGACAGAGCTGTGACCCACCACTTCCGCTACTAGCTA 367
Db 993 CCATCTGCCGACAGAGCTGTGACCACTTCCGCTACTAGCTA 1036

RESULT 2
AC136382
LOCUS AC136382
DEFINITION Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
AC136382
AC136382.1 GI:244662257
VERSION HTG, HTGS_PHASE1
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 185148)
AUTHORS Muzny,D,Matie, Metzker,M,Lee, Abranzon,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsdorfs,S, Amin,A, Anguiano,D, Anyalebechi,V, Ayagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guayana,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hledun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howells,S, Huylk,S, Hume,U, Idledit,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lotenshewa,L, Louised,H, Lozado,R,U, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhinney,S, McLeod,M, McNeill,T, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwokelelelele,O, Okononwu,G, Olanunsgoon,A, Pal,S, Parks,K, Paeterak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Polidexter,A, Popovic,D, Primus,E, Pu,L, Puazo,M, Quiroz,J, Rachin,E, Reeves,K, Regier,M,A, Reigh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojka,A, Rose,M, Rose,R, Ruiz,S,U, Sanders,W, Saverly,G, Scheer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartbeyan,A, Sisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,R, Sosa,J, Steimle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Uemari,K, Valdes,R, Vera,V, Villasea,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Wilson,R, Wleczek,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
Rat Genome Sequencing Consortium.
Unpublished
2 (bases 1 to 185148)

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Center project name: KDGS
Center clone name: CH230-97018
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 1000 of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 135613 bases at least Q40
Consensus quality: 140849 bases at least Q30
Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 63 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1010: contig of 1010 bp in length
* 1011 1110: gap of unknown length
* 1111 2608: contig of 1498 bp in length
* 2609 2708: gap of unknown length
* 2709 3915: contig of 1207 bp in length
* 3916 4015: gap of unknown length
* 4016 5547: contig of 1532 bp in length
* 5548 5647: gap of unknown length
* 5648 7119: contig of 1472 bp in length
* 7120 7219: gap of unknown length
* 7220 8885: contig of 1666 bp in length
* 8886 8985: gap of unknown length
* 8986 10258: contig of 1273 bp in length
* 10259 10359: gap of unknown length
* 10359 11626: contig of 1268 bp in length
* 11627 11726: gap of unknown length
* 11727 13488: contig of 1762 bp in length
* 13489 13588: gap of unknown length
* 13589 15122: contig of 1534 bp in length
* 15123 15223: gap of unknown length
* 15223 15945: contig of 1724 bp in length
* 15946 16947: gap of unknown length
* 16947 17046: gap of unknown length
* 17047 18164: contig of 1118 bp in length
* 18165 18264: gap of unknown length
* 18265 19678: contig of 1414 bp in length
* 19679 19779: gap of unknown length
* 19779 21687: contig of 1909 bp in length
* 21688 21787: gap of unknown length
* 21788 22892: contig of 1105 bp in length
* 22893 22993: gap of unknown length
* 22993 24336: contig of 1344 bp in length
* 24337 24437: gap of unknown length
* 24437 26617: contig of 2181 bp in length
* 26618 26717: gap of unknown length
* 26718 28577: contig of 1860 bp in length
* 28578 28677: gap of unknown length
* 28678 30823: contig of 2146 bp in length
* 30824 30923: gap of unknown length
* 30924 32265: contig of 1343 bp in length
* 32266 32366: gap of unknown length
* 32367 33894: contig of 1528 bp in length
* 33895 33994: gap of unknown length
* 33995 35373: contig of 1379 bp in length

```

* 35473: gap of unknown length
* 35474: gap of unknown length
* 37295: contig of 1822 bp in length
* 37296: gap of unknown length
* 37395: contig of 1025 bp in length
* 37396: gap of unknown length
* 38421: contig of 1454 bp in length
* 38521: gap of unknown length
* 39975: contig of 2209 bp in length
* 40074: gap of unknown length
* 42283: contig of 2209 bp in length
* 42284: gap of unknown length
* 43721: contig of 1338 bp in length
* 43722: gap of unknown length
* 43822: contig of 2717 bp in length
* 46538: gap of unknown length
* 46639: contig of 1983 bp in length
* 48622: gap of unknown length
* 48721: contig of 1881 bp in length
* 48722: gap of unknown length
* 50603: contig of 3139 bp in length
* 50703: gap of unknown length
* 53841: contig of 2254 bp in length
* 53842: gap of unknown length
* 56195: contig of 2654 bp in length
* 56196: gap of unknown length
* 56295: contig of 2654 bp in length
* 58949: gap of unknown length
* 59049: gap of unknown length
* 61032: contig of 1983 bp in length
* 61033: gap of unknown length
* 61133: contig of 1810 bp in length
* 62942: gap of unknown length
* 62943: gap of unknown length
* 63043: contig of 3003 bp in length
* 66045: gap of unknown length
* 66145: gap of unknown length
* 66146: contig of 2052 bp in length
* 68197: gap of unknown length
* 68297: gap of unknown length
* 68298: contig of 3938 bp in length
* 72235: gap of unknown length
* 72236: gap of unknown length
* 72336: contig of 3468 bp in length
* 75803: gap of unknown length
* 75804: gap of unknown length
* 75904: contig of 2452 bp in length
* 78355: gap of unknown length
* 78356: gap of unknown length
* 78455: contig of 3003 bp in length
* 79455: gap of unknown length
* 81455: contig of 3115 bp in length
* 81459: gap of unknown length
* 84673: contig of 3115 bp in length
* 84674: gap of unknown length
* 84773: contig of 2966 bp in length
* 87739: gap of unknown length
* 87740: gap of unknown length
* 87840: contig of 3636 bp in length
* 91535: gap of unknown length
* 91536: gap of unknown length
* 91635: contig of 3887 bp in length
* 91636: gap of unknown length
* 95622: contig of 3108 bp in length
* 95623: gap of unknown length
* 96731: gap of unknown length
* 98830: contig of 2788 bp in length
* 98831: gap of unknown length
* 101618: contig of 3817 bp in length
* 101619: gap of unknown length
* 105535: contig of 3758 bp in length
* 105536: gap of unknown length
* 109393: contig of 4071 bp in length
* 109394: gap of unknown length
* 109493: contig of 3966 bp in length
* 11365: gap of unknown length
* 11366: gap of unknown length
* 117631: contig of 4772 bp in length
* 117631: gap of unknown length
* 117731: contig of 4042 bp in length
* 122503: gap of unknown length
* 122503: contig of 4042 bp in length
* 12644: gap of unknown length
* 12645: gap of unknown length

```

Query Match 57.6%; Score 214.4; DB 2; Length 185148;
 Best Local Similarity 97.3%; Pred. No. 1.9e-51;
 Matches 218; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 144 TCGCCCATGTCAAGCAAACTTATTTTGAACGCGGATCTTAGACGCTGCC 203
 DB 99243 TGTGCGAGCTCAAGCAAACTTATTTTGAACGCGGATCTTAGACGCTGCC 99302
 QY 204 CTGACATCATTAACCGTGTGCGAGCGACCCCTTATTAAGCGCCCTGGTATGCGCAG 263

```

DB 99303 CTGACATCATTAACCGTGTGCGAGCGACCCCTTATTAAGCGCCCTGGTATGCGCAG 99362
QY 264 CCAGCATGTCTCACTGCCCCCGCGAGACCAAAACCAGGAGACTTGAACACTGCACACG 323
DB 99363 CCAGCATGTCTCACTGCCCCCGCGAGACCAAAACCAGGAGACTTGAACACTGCACACG 99422
QY 324 CCATGTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
DB 99423 CCATGTGCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 99466

RESULT 3
MMLIF
LOCUS Mmusculus 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493.1 GI:52699
VERSION X67493.1 GI:52699
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
JOURNAL Hepatology 19 (3), 656-665 (1994)
MEDLINE 94164648
PUBMED 7509771
REFERENCE
AUTHORS Mohn,K.L., Waddell,J.R. and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
JOURNAL Nucleic Acids Res.
REFERENCE 3 (bases 1 to 1363)
AUTHORS Taub,R.A.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curtie Boulevard, Philadelphia, PA 19104-6145, USA

FEATURES
source
1..1363
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="UM2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
/note="AP-2 consensus site"
692..706
/note="insulin-responsive element"
720..741
/note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
792..1336
/number=1
792
/note="mRNA cap site"
964..>1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"

```

/translation="MPEFLTVSWPFLILLSFOIGVAAQAPQWHCAPCTAERLGLCP
PVPASCPESIRPAGCGCCPTCALPMGAACGVATARCAQGISCRALPGEPRPLHALTRG
QGASLPEPAPATSTLFSQHE"

Query Match	50.8%	Score 188.8	DB 10	Length 1363
Best Local Similarity	90.2%	Pred. No. 4.3e-44		
Matches 202; Conservative	0	Mismatches 22	Indels 0	Gaps 0

QY	144	GGCGCCCATGTCACAGCAAGAAAAACAACCTTATTTTGAACACGGGGATCCTAGACAGCTGGC	203
Db	667	TGTGTAGAGCTCACAGCAAGAAAACAAACCTTATTTTGAACACTGGGGTCTTAGACAGCTGGC	726
QY	204	CTGACAAATCATTTAAACCCGTCGTGCAGGCACGCCCTTACATAAGGCGCTGGATGAGCCAG	263
Db	727	CTGACAAATCATTTAACTGTGTGCCGCACAGCAGCCTTCAATAAGGCTTGGGTATGACACAG	786
QY	264	CCAGCATAGTCCACTGCGCCGCCGAGACACAAACCGACGAGCATTTGAACATCTGCACACGG	323
Db	787	CCAGCATAGTCCACTGCGCCGCCGAGACACACACCGACGAGCATTTGAACATCTGCACACGG	846
QY	324	CCATTTGCCACAGAGCTGTGACCAACCACTTCCGTACTAGCTA	367
Db	847	CCGTTGCCACAGAGCTGTGACCAACCACTTGCACACTACTATTTA	890

RESULT 4	
LOCUS	AL607124
DEFINITION	AL607124 192943 bp DNA linear ROD 11-APR-2002 Mouse sequence from clone RP23-20C9 on chromosome 11, complete sequence.
ACCESSION	AL607124
VERSION	AL607124.15 GI:20145926
KEYWORDS	HTG.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1
REFERENCE	Oliver, K.
AUTHORS	Direct Submission
TITLE	Submitted (11-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL	Cambri@geshare, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Apr 12, 2002 this sequence version replaced gi 19847866.
COMMENT	

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the features table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP. Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-20C9 is from the RPc1-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: DBAC3.6.

```
FEATURES
  source      location/qualifiers
1. 192843    1. 192843
           /organism="Mus musculus"
           /mol_type="genomic DNA"
           /db_xref="taxon:10090"
           /chromosome="11"
```

```

/c1one="RP23-20C9"
/c1one_lib="RPCI-23"

```

Query Match	50.8%;	Score 188.8;	DB 10;	Length 192843;
Best Local Similarity	90.2%;	Pred. No. 5.5e-44;		
Matches 202;	Conservative	0;	Mismatches 22;	Indels 0;
				Gaps 0;

QY	144	TTGGCGCCATGTGCACAGAGAAACAAACTTATTTTGAACACGGGGATCCTACACGCTGCC	203
Db	2419	TTGTGAGGCTGCACAGAGAAACAAACTTATTTTGAACACGGGGTCTTACACGCTGCC	2478
QY	204	CTGACAAATCATTTAAACCCGTGCTGCAGACACGCCCTTCATTAAGGCCCTCGGTTATGGCCAG	263
Db	2479	CTGACAAATCATTTAAACCTGTGTGCCGACAGCCGCTTCATTAAGGCTCTGGGTTAGCACAG	2538
QY	264	CCAGCATGTGTCAACTGCCCGCCGAGACACAAACCCAGCGAGCATTTGAACACTGCACACGG	323
Db	2539	CCAGCATGTGTCAACTGCCCGCGGAGACACACCCAGCGAGCATTTGAACACTGCACACGG	2598
QY	324	CCATCTGCCCGAGAGCTGTGACACACACTTCGGCTACTAGCTA	367
Db	2599	CCGCTGTGCCCGAGAGCTGTGACACACACTTCGCACTACTACTTA	2642

RESULT 5	BC013345	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
	BC013345	1569 bp	mRNA linear	BC013345	BC013345	GI:15426482	Mus musculus (house mouse)	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1569)	Struhsberg, R. L., Feingold, E. A., Grouse, J. H., Derge, J. G.,

JOURNAL PUBMED REFERENCE AUTHORS TITLE JOURNAL	Almouzni, R.D., Collins, F.S., Wagner, L., Schmeum, C.M., Schaefer, C.F., Shat, N.K., Altschul, S.F., Zeeberg, B., Buelow, K.H., Schaefer, C.F., Shat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Ditchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stetlton, M., Soares, M.B., Donald, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Ustin, T.B., Tschibiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Milahy, S.J., Bosak, S.A., McEwan, P.J., McEwan, R.V., Malek, U.A., Gunaratne, P.H., Richards, S., Weller, D.C., Hale, S., Garcia, A.M., Gay, L.J., Huily, S.W., Viallalon, K.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shchavchenko, Y., Bouffard, G.C., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schermer, A., Schein, J.E., Jones, S.U. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002) 12477932 2 (bases 1 to 1569) Strausberg, R. Direct Submission Submitted (31-AUG-2001) National Institutes of Health, Mammalian
---	--

REMARK	COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov	
Contact: MGC help desk	
Email: cgabs-r@mail.nih.gov	
Tissue Procurement: Jeffrey E. Green, M.D.	
cDNA Library Preparation: Life Technologies, Inc.	
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)	

DNA Sequencing by: Institute for Systems Biology
<http://www.systembiology.org>
 Contact: amadan@systembiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MQC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMN at: <http://image.llnl.gov>
 Series: IRK Plate: 18 Row: 1 Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
 Location/Qualifiers

FEATURES

source

1.1569
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_mgc="MGCL14075 IMAGE:4161889"
 /tissue_type="liver, normal. 5 month old male mouse."
 /clone_id="NCI_CGAP_L19"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1.1569
 /gene="Igf1bp1"
 /note="synonym: IGFBP-1"
 /db_xref="LocusID:16006"
 /db_xref="MGI:96436"
 204.1022
 /gene="Igf1bp1"
 /codon_start=1
 /product="Insulin-like growth factor binding protein 1"
 /protein_id="AAH13345.1"
 /db_xref="GI:15426483"
 /db_xref="LocusID:16006"
 /db_xref="MGI:96436"
 /translation="MPEPLTVSWPRLILSFQIVAGAPQPMHCAPCAETARLGLCP
 PVPASCPRLSRPAGCCCTCALPMGACGVATARGASLRALPGEPRPLHALTRG
 QGACVPEPATPTASGLSIEKEAKASVPRVPESEKMTBEOQLDSFHLMAFSD
 OPILMNAISTYKSMRAREIMADIKWKQPCRELIVLERLAAQKAGEEIKFYLPN
 CNKNGFYHSKQCEITLSDGEARLQCMCVPMGSRRIPLSLERDPRCHQFYVNVN"

CDS

gene

ORIGIN

Query Match 26.4%; Score 98.2; DB 10; Length 1569;
 Best Local Similarity 92.8%; Pred. No. 1.2e-17;
 Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 257 TGGCCAGCAGCATGCTCCAGTCCCGCCGAGAGACAAACGAGGAGCATGGAACACTG 316
 |||
 DB 20 TGGGACCCAGCATGTGCTCCAGTCCCGCCGAGAGACACACCGAGCATGGAACACTG 79
 |||
 QY 317 CACACGGCCATCTGCCAGAGAGCTGTGACCAACCTTCGCTACTAGCTA 367
 |||
 DB 80 CACACGGCCGTCTGCCAGAGAGCTGTGACCAACCTTCGCTACTATCTA 130
 |||

RESULT 6
 LOCUS AY560836 1507 bp mRNA linear ROD 22-MAR-2004
 DEFINITION Spermophilus tridecemlineatus insulin-like growth factor binding
 protein 1 (IGFBP1) mRNA, complete cde.
 ACCESSION AY560836
 VERSION AY560836.1 GI:45505308
 KEYWORDS

SOURCE

Spermophilus tridecemlineatus (thirteen-lined ground squirrel)

Spermophilus tridecemlineatus

Mammalia; Eutheria; Rodentia; Scuriognathi; Scuriidae; Scuriinae;
 Spermophilus.

1 (baaes 1 to 1507)

Li, Y., Klimants, D. and Hallenbeck, J.M.

Cloning and characterization of insulin-like growth factor binding
 protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL Unpublished
 REFERENCE 2 (baaes 1 to 1507)
 AUTHORS Li, Y., Klimants, D. and Hallenbeck, J.M.
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
 Bethesda, MD 20892, USA

FEATURES

source

1.1507
 /organism="Spermophilus tridecemlineatus"
 /mol_type="mRNA"
 /db_xref="taxon:43179"
 1.1507
 /gene="IGFBP1"
 689.1507
 /gene="IGFBP1"
 /codon_start=1
 /product="Insulin-like growth factor binding protein 1"
 /protein_id="AAS67029.1"
 /db_xref="GI:45505308"
 /translation="MPEVPAAGLMPRLILAVQVSTVASTOPMHCAPCSAKLALCP
 FVSSCPRLSRPAGCCPCALPLGACGVATARGASLRALPGEPRPLHALTRG
 QGACVPEPATPTASGLSIEKEAKASVPRVPESEKMTBEOQLDSFHLMAFSD
 OPILMNAISTYKSMRAREIMADIKWKQPCRELIVLERLAAQKAGEEIKFYLPN
 CNKNGFYHSKQCEITLSDGEARLQCMCVPMGSRRIPLSLERDPRCHQFYVNVN"

CDS

ORIGIN

Query Match 21.3%; Score 79.2; DB 10; Length 1507;
 Best Local Similarity 67.1%; Pred. No. 4e-12;
 Matches 149; Conservative 0; Mismatches 58; Indels 15; Gaps 2;

QY 144 TGGCCCATGTGCACAGAAACAACTATTGGAACAGCGGATCTTAGACGCTGCC 203
 |||
 DB 420 TGGGAGAAAGCAGAGGAAACAACTATTGGAACACGCGCTCTTAGACGCGCG 479
 |||
 QY 204 CTGACATCATTTAA-----CCCGTGTCCGAGCCAGCCCTTCATAGGCGCTTG 253
 |||
 DB 480 CTGGCAATCATTTACTTTCTAGTCCAGTGAAGTAACCGGCCCTTTATAGGCAAGG 539
 |||
 QY 254 GTATGGCCAGCCAGCATGTGCACCTGCCCGCCGAGACAAACCGAGGAGCATTTGAACA 313
 |||
 DB 540 CTGCACCAGCAGAGCATGTGCACCTGCCCGGAGACTCAGAAATTGACACTATCA-- 597
 |||
 QY 314 CTGCACAGCCCATCTGCCCCAGAGAGCTGTGACCAACCACTTC 355
 |||
 DB 598 ---GCCACTGTCCAGCTGCCAGAGAGCATGACCACTGTCC 636
 |||

RESULT 7
 LOCUS AX401932 1500 bp DNA linear PAT 06-JUN-2002
 DEFINITION Sequence 1608 from Patent WO0210453.
 ACCESSION AX401932
 VERSION AX401932.1 GI:21338112
 KEYWORDS

SOURCE

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae;
 Rattus.

Mendrick, D., Porter, M.W., Johnson, K.R., Casale, A.L. and

Elschoff, M.R.

Molecular toxicology modeling

Patent: WO 0210453-A 1608 07-FEB-2002;

Gene Logic, Inc. (US)

location/Qualifiers

1.1500

/organism="Rattus norvegicus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:10116"
 /note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

```

Query Match          20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 349
      |||
DB 8 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 8
AX827271 1500 bp DNA linear PAT 12-DEC-2003
LOCUS Sequence 5 from Patent EP1344834.
ACCESSION AX827271
VERSION AX827271.1 GI:39837360
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Boess, F., Suter-Dick, L. and Wolf, D.
TITLE Methods for the toxicity prediction of a compound
JOURNAL Patent: EP 1344834-A 5 17-SEP-2003;
          F. HOFFMANN-LA ROCHE AG (CH)
FEATURES
source 1..1500
        /organism="Rattus norvegicus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:10116"

ORIGIN
Query Match          20.5%; Score 76.4; DB 6; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 349
      |||
DB 8 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 9
RATIGFB 1500 bp mRNA linear ROD 27-APR-1993
DEFINITION Rat IGF binding protein-1 (RIGFBP-1) mRNA, complete cds.
ACCESSION M58634
VERSION M58634.1 GI:204732
KEYWORDS IGF binding protein-1.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Mohn, K.L., Melby, A.E., Tewari, D.S., Laz, T.M. and Taub, R.
TITLE The gene encoding rat insulinlike growth factor-binding protein 1
JOURNAL Mol. Cell. Biol. 11 (3), 1393-1401 (1991)
MEDLINE 91141487
PUBMED 1705004
COMMENT Original source text: Rat, cDNA to mRNA.
FEATURES
source 1..1500
        location/Qualifiers
        /organism="Rattus norvegicus"

```

```

/mol_type="mRNA"
/strain="Fish"
/db_xref="taxon:10116"
/tissue_type="regenerating liver"
/dev_stage="adult"
160..978
/codon_start=1
/product="IGF binding protein-1"
/protein_id="AAA41380.1"
/db_xref="GI:204733"
/translation="MPEFLTVVSPFILLISFQRYVYAGAPQPHCAPCTARELECP
PVPAACEIPSPAGCGCCPTCALPLGACGATNCAQGLSCRLPEPPLHLATG
QGACVLEPAAPATSLSGSOHEBAKAAVASDELVAESPEMTEQLDLSFILMAPSRD
QPILMNAISTVSSMRAREITDILKKWKEPCORELTVBERLAAAOAKAGDEIYKFLPN
CNKNGFYHSKQCETSLDGBALGCMCVYPMWGSKI PGLSLRTGDPNCHQYFNQVQ"
160..234
235..975
/product="IGF binding protein-1"

ORIGIN
Query Match          20.5%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 2.6e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 349
      |||
DB 8 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 67

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 68 CACTTCGGCTACTATCTA 85

RESULT 10
AX163782 5001 bp DNA linear PAT 22-JUN-2001
LOCUS Sequence 46 from Patent WO0138579.
ACCESSION AX163782
VERSION AX163782.1 GI:14544878
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE
AUTHORS Gould-Rothberg, B.E., Dipippo, V.A., Rameeh, T.M. and Gerwein, R.W.
TITLE Method of identifying toxic agents using msalD-induced differential
JOURNAL gene expression in liver
PUBMED Patent: WO 0138579-A 46 31-MAY-2001;
          Curagen Corporation (US)
FEATURES
source 1..5001
        location/Qualifiers
        /organism="Rattus norvegicus"
        /mol_type="unassigned DNA"
        /db_xref="taxon:10116"

ORIGIN
Query Match          20.5%; Score 76.4; DB 6; Length 5001;
Best Local Similarity 98.7%; Pred. No. 2.8e-11;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 290 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 349
      |||
DB 1 CACAAACCCGAGCAGTGAACACTGCACACGGCCATCTGCCGAGAGCTGTGACAC 60

QY 350 CACTTCGGCTACTAGCTA 367
      |||
DB 61 CACTTCGGCTACTATCTA 78

RESULT 11
RATIGFB

```

LOCUS RATTIGBA 5001 bp DNA linear ROD 30-NOV-1995
 DEFINITION Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,
 complete cds.
 ACCESSION L22979
 VERSION L22979.1 GI:1098472
 KEYWORDS insulin-like growth factor binding protein-1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (baaes 1 to 5001)
 AUTHORS Iacson, R., Oehler, D., Yang, E., Goswami, R. and Unterman, T.
 TITLE Bideoxy sequencing and structural analysis of the rat insulin-like
 growth factor binding protein-1 gene
 JOURNAL Biochim. Biophys. Acta 1218 (1), 95-98 (1994)
 MEDLINE 94250701
 PUBMED 7514892
 COMMENT On Nov 30, 1995 this sequence version replaced gi:385167.
 Original source text: Rattus norvegicus (strain Sprague-Dawley)
 DNA.

FEATURES
 source Location/Qualifiers
 1..5001
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /sex="male"
 /cissue_type="liver"
 /dev_stage="adult"
 /cissue_1fb="Clontech EMBL-3 Sp6/T7"
 60..525
 /number=1
 exon 60..152
 153..4221
 /gene="IGFBP-1"
 join(153..525,1850..2034,2874..3002,4090..4221)
 /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein"
 /protein_id="AAA82581.1"
 /db_xref="GI:1098473"
 /translation="MPELVTVSWPELILSRQVNVVAGAPQPHHCAPCTARLELCP
 EYPASCEISRPAGCGCCCTCALPLGAAAGVATAACQGLSCRALGEPRLHALITRG
 OGACVLEPAAPATSSLSGSQHEAAVAASEDELSPEMTESQLDSEPHLMAPERED
 QPILMNAISTVSMRAREITDDKKWKEPCORLIVLESLAAQCKAGGEIKFYLPLN
 CNKNGFYHSKOCETSLDGBAGLCWCVCYPMGSKIKPSLETRDPPNCHQYFNQV"
 526..1849
 /gene="IGFBP-1"
 /number=1
 exon 1850..2034
 /gene="IGFBP-1"
 /number=2
 2035..2873
 /gene="IGFBP-1"
 /number=2
 2874..3002
 /gene="IGFBP-1"
 /number=3
 3003..4089
 /gene="IGFBP-1"
 /number=3
 4090..4743
 /number=4
 4222..4743
 exon
 3'UTR
 ORIGIN
 Query Match 20.5%; Score 76.4; DB 10; Length 5001;
 Best Local Similarity 98.7%; Pred. No. 2,8e-11;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

290 CACAAACCGAGCATGTGAACACTGCACAGCGCATGTGCCAGAGAGCTGTGACAC 349
 |||||

Db 1 CACAAACCGAGCATGTGAACACTGCACAGCGCATGTGCCAGAGAGCTGTGACAC 60
 QY 350 CACTTCCGCTACTACTA 367
 Db 61 CACTTCCGCTACTACTA 78

RESULT 12
 BC078889 1510 bp mRNA linear ROD 03-AUG-2004
 LOCUS Rattus norvegicus insulin-like growth factor binding protein 1,
 mRNA (cDNA clone MGC:93595 IMAGE:1129185), complete cds.
 ACCESSION BC078889
 VERSION BC078889.1 GI:50927646
 KEYWORDS MGC.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (baaes 1 to 1510)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Umed, T.B., Toshiyuki, S.,
 Carrinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalka, U., Smalls, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (baaes 1 to 1510)
 DIRECTOR Director MGC Project.
 REFERENCE Submitted (02-AUG-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdelpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 184 Row: e Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6981079.
 Location/Qualifiers
 1..1510
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"

FEATURES
 source

gene

cds

ORIGIN

Query Match
Best Local Similarity 98.5%; Pred. No. 4,3e-08; Length 1510;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 301 CGAGCATTGAACACTGCACACGGCCATGTGCCACAGAGCTGTGACCAACCACTTCCGCTA 360
Db 1 CGAGCATTGAACACTGCACACGGCCATGTGCCACAGAGCTGTGACCAACCACTTCCGCTA 60

Qy 361 CTAGCTA 367
Db 61 CTATCTA 67

RESUT.T 13
LOCUS AY095345
DEFINITION Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
ACCESSION AY095345
VERSION AY095345
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 3886)
Kim,J.-J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Utemman,T.G.
Regulation of insulin-like growth factor binding protein-1 promoter
activity by FKHR and HOXA10 in primate endometrial cells
Biol. Reprod. 68 (1), 24-30 (2003)
12493691
2 (bases 1 to 3886)
Kim,J.-J., Jaffe,R.C. and Fazleabas,A.T.
Direct Submission
Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source

gene

Location/Qualifiers
1..3886
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
3655..>3886
/gene="IGFBP-1"

mRNA
 <3655..>3886
 /gene="IGFBP-1"
 /product="insulin-like growth factor binding protein-1"
 3655..>3886
 /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein-1"
 /protein_id="AA023273.1"
 /db_xref="GI:20853765"
 /translation="MSEVPFARVWLVELLTVGVVTAAPWOCAPCSAEKTLCPVPV"
 PASCEVTRSAAGCCCPKCALPLGAAGVATAPR"

ORIGIN
 Query Match 17.2%; Score 64; DB 9; Length 3886;
 Best Local Similarity 60.6%; Pred. No. 1.1e-07;
 Matches 149; Conservative 0; Mismatches 80; Indels 17; Gaps 2;

CDS
 Db 120 TCCAGGAACCAAGGAGTGGCCCGTGGCCCATGTCAACAAGCAAAACAATTATTTTGA 179
 3331 TCCTCCCACTCGCGGTTTGGCTTGAAGGCGCTTGGCGCACTAGCAAAACAATTATTTTGA 3390
 Oy ACACGGGGATCTTAGACAGCTGCGCTGACATCATTAATTAACC-----GTGCTGCCG 229
 Db 3391 ACATCTAGACTCTTAGCGCGCGCGCTGCCATCATTAACCTCTGTGCAAGTGGCGCG 3450
 Oy 220 AGCCAGCCCTTCATPAAGGCGCTGGGTAATGCGACCCAGACATGTCCACTGCGCGCGAGA 289
 Db 3451 CCTGTGCGCTTATTAAGGCGCGCGCTGTGTCCAGCAAGCATGGCGCACCGCATCCATC 3510
 Oy 290 CACAAACCCAGCAGCATTTGAACACTGACACCGGCCATCTGCCAGAGAGTGTGACAC 349
 Db 3511 CAGCAAGC-----ATCTGCCGCGCGCGCGCGCCACCTCTCCAGAGACATGGCCAC 3563
 Oy 350 CACTTC 355
 Db 3564 CGCTCC 3569

RESULT 14
 AX409747 6128 bp DNA linear PAT 14--JUN-2002
 LOCUS
 DEFINITION Sequence 2394 from Patent WO0229103.
 AC409747
 VERSION AX409747.1 GI:21442452
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1
 Alvaras, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
 Gene expression profiles in liver cancer
 Patent: WO 0229103-A 2394 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source
 1..6128
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. M74587"

ORIGIN
 Query Match 17.2%; Score 63.8; DB 6; Length 6128;
 Best Local Similarity 63.2%; Pred. No. 1.3e-07;
 Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

Db 120 TCCAGGAACCAAGGAGTGGCCCGTGGCCCATGTCAACAAGCAAAACAATTATTTTGA 179
 438 TCCTCCCACTCGCGGTTTGGCTTGAAGGCGCTTGGCGCACTAGCAAAACAATTATTTTGA 497
 Oy 180 ACACGGGGATCTTAGACAGCTGCGCTGACATCATTAATTAACC-----GTGCTGCCG 229
 Db 498 ACATCTAGACTCTTAGCGCGCGCGCTGCCATCATTAACCTCTGTGCAAGTGGCGCG 557

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 2828.19 Seconds
(without alignments)
5006.706 Million cell updates/sec

Title: US-09-972-916B-5

Perfect score: 372

Sequence: 1 tacactggggggccagagatcc.....ttcgcctactagctagccgc 372

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_esc1:*

2: gb_esc2:*

3: gb_hic:*

4: gb_esc3:*

5: gb_esc4:*

6: gb_esc5:*

7: gb_esc6:*

8: gb_gsa1:*

9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.2	26.4	1013	2	BF236974
2	87.8	23.6	310	6	BY783538
3	87.8	23.6	310	6	BY783538
4	87.8	23.6	354	6	BY794229
5	86.8	23.3	401	1	AI785818
6	86.8	23.3	480	1	AI196314
7	86.8	23.3	706	1	AI196314
8	86.8	23.3	765	1	AI196314
9	86.8	23.3	785	1	AI196314
10	86.8	23.3	799	1	AI196314
11	86.8	23.3	811	6	CA478518
12	86.8	23.3	848	1	AI790802
13	86.8	23.3	852	1	AI1528304
14	85.2	22.9	605	1	AI196154
15	82.2	22.1	380	1	AI1785039
16	81.4	21.9	846	7	CO573026
17	80.4	21.6	692	7	CV127049
18	79.4	21.3	713	7	CO560662
19	75.4	20.3	748	7	CO575629
20	73.8	19.8	694	7	CV117001
21	72.4	19.7	632	7	W30013
22	72.4	19.5	618	6	BB660958
23	72.4	19.5	618	6	CD561711
24	72.4	19.5	632	1	AI892189

25	67.4	18.1	488	1	AA674302
26	65.4	17.6	801	7	CK472246
27	64.4	17.3	840	7	CK473709
28	62.8	16.9	759	1	AA105355
29	61.8	16.6	269	2	BB604790
30	61.6	16.6	587	2	AW916227
31	60.8	16.3	1589	3	CR621807
32	60.8	16.3	1601	3	CR595377
33	59	15.9	545	1	AA060360
34	47.8	12.8	1159	8	CNS015XR
35	47.2	12.7	659	9	AZ840793
36	44.4	11.9	1128	3	CR657703
37	44.4	11.9	1128	3	CR671810
38	44.4	11.9	1159	3	CR663312
39	44.4	11.9	1167	3	CR667848
40	42.2	11.3	597	8	AZ652514
41	41.6	11.2	325	2	AW855818
42	40	10.8	212	2	BR149547
43	39.4	10.6	500	8	AO612859
44	39.2	10.5	234	2	AW886850
45	39.2	10.5	289	2	BE066031

ALIGNMENTS

RESULT 1
LOCUS BF236974
DEFINITION 602026605F1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:416189 5',
LOCUS BF236974
ACCESSION BF236974
VERSION BF236974.1 GI:11150891
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM9443 row: j column: 02
High quality sequence at: 581.
Location/Qualifiers
1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:416189"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Query Match 26.4%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 5.9e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
CY 257 TGGCAGCAGCAGTGGTCCCGCCGAGACAAACAGCAGCAGTGAACACTG 316

[illegible]

DB		62	TGCCAAGAGACTGTGACCACTATTGCACCTACTATCTA	100
RESULT_3				
BY794229				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
AUTHORS				
REFERENCE				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
COMMENT				
FEATURES				
source				
ORIGIN				
Query Match				
Best local Similarity				
Matches				
OY				
DB				
OY				
DB				
RESULT_4				
BY771317				

LOCUS BY77317 354 bp mRNA linear EST 23-MAR-2004
DEFINITION BY77317 RIKEN full-length enriched, 17.5 days embryo whole body
MUS musculus cDNA clone L930052G15 5', mRNA sequence.
ACCESSION BY77317
VERSION BY77317.1 GI:39697955
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugawara, Y., Saito, R., Oosato, N., Fukuda, S., Sato, K., Matshiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gueirincich, S., Beisel, K., Pavan, W., Aldinis, V., Nakagawa, A., Held, W. A., Iwata, H., Kono, T., Nakaschi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Hensch, T. K., Brinkmeyer, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
TITLE Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia
JOURNAL Genome Res. 13 (6B), 1273-1289 (2003)
MEDLINE 22703353
PUBMED 12819125
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.jp/) for further details.
FEATURES
source Location/Qualifiers
1..354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930052G15"
/issue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_1fb="RIKEN full-length enriched, 17.5 days embryo whole body"
ORIGIN
Query Match 23.6%; Score 87.8; DB 6; Length 354;
Best Local Similarity 92.9%; Pred. No. 4.3e-14;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 269 ATGTCTCACTGCCCCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 328
DB 2 ATGTCTCACTGCCCCGCGAGACACACACCCAGGAGCATTTGAACACTGACACGGGCATC 61
QY 329 TGCCGAGAGGCTGTGACCAACCACTTCCGCTACTACTA 367
DB 62 TGCCGAGAGGCTGTGACCAACCACTTCCGCTACTACTA 100
RESULT 5
LOCUS A1785818 401 bp mRNA linear EST 02-UTL-1999
DEFINITION u178h05.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1885859 5' similar to gb:XB1579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
ACCESSION A1785818

VERSION A1785818.1 GI:5333534
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rutter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The MASHU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Other_ESTS: u178h05.x1
Contact: Marra M./MASHU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1600
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972893
Seq primer: custom primer used
High quality sequence scop: 126.
FEATURES
source Location/Qualifiers
1..401
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1885859"
/sex="female"
/dev_stage="adult"
/lab_host="DHL0B"
/clone_1fb="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME185-F13; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGCGCTTCTG]; digested and cloned into distinct DraIII sites of the pME185-F13 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTGTCTTAAGCTGG and 3' end primer CGACTGCACTCGACGACA."
ORIGIN
Query Match 23.3%; Score 86.8; DB 1; Length 401;
Best Local Similarity 92.9%; Pred. No. 8.5e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 270 TGTCTCACTGCCCCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 329
DB 1 TGTCTCACTGCCCCGCGAGACACACACCCAGGAGCATTTGAACACTGACACGGGCATC 60
QY 330 GCCCAGAGAGGCTGTGACCAACCACTTCCGCTACTACTA 367
DB 61 GCCCAGAGAGGCTGTGACCAACCACTTCCGCTACTACTA 98
RESULT 6
LOCUS A1196314 480 bp mRNA linear EST 14-OCT-1998
DEFINITION u171a07.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1887828 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:XB1579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 480)
 REFERENCE
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:972152
 FEATURES
 source High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-F13; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCTTACTGG), digested and cloned into distinct DraIII sites of the pME18S-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGAGCTCGACACA."
 ORIGIN
 Query Match 23.3%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 8.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 270 TGTTCACATGCCCCGAGACACAAACCCAGGAGCATTTGAACACTGACACAGGCGCATCT 329
 Db 1 TGTTCACATGCCCCGAGAGACACACACCCAGGAGCATTTGAACACTGACACAGGCGCTT 60
 QY 330 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 367
 Db 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
 RESULT 7
 LOCUS A1530146 706 bp mRNA linear EST 18-MAR-1999
 DEFINITION u189f09.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:W5316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146 GI:4444281
 VERSION A1530146.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 706)
 REFERENCE
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 FEATURES
 source High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-F13; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor (TGTGGCTTACTGG), digested and cloned into distinct DraIII sites of the pME18S-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTCGAGCTCGACACA."
 ORIGIN
 Query Match 23.3%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 9.5e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 270 TGTTCACATGCCCCGAGACACAAACCCAGGAGCATTTGAACACTGACACAGGCGCATCT 329
 Db 1 TGTTCACATGCCCCGAGAGACACACACCCAGGAGCATTTGAACACTGACACAGGCGCTT 60
 QY 330 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 367
 Db 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
 RESULT 8
 LOCUS A1098594 765 bp mRNA linear EST 20-AUG-1998
 DEFINITION u21e07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Scheinberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 395.
Location/Qualifiers
1..765
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1481988"
/sex="female"
/dev_stage="adult"
/lab_host="DHI0B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTCTCTAAAGCTGCG and 3' end primer CGACTGCACTGAGCACA."

ORIGIN
Query Match 23.3%; Score 86.8; DB 1; Length 765;
Best Local Similarity 92.9%; Pred. No. 9.7e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCACTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 329
DB 1 TGGTCACTGCCCGCGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 60

QY 330 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 98

RESULT 9
LOCUS A1530313 785 bp mRNA linear EST 18-MAR-1999
DEFINITION u167c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313 GI:4444448
VERSION A1530313.1 GI:4444448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974109
Seq primer: custom primer used
High quality sequence stop: 459.
Location/Qualifiers
1..785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DHI0B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FU3; Site_1: DraIII (CACTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTGG]; digested and cloned into distinct DraIII sites of the pME18S-FU3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTCTCTAAAGCTGCG and 3' end primer CGACTGCACTGAGCACA."

ORIGIN
Query Match 23.3%; Score 86.8; DB 1; Length 785;
Best Local Similarity 92.9%; Pred. No. 9.7e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 270 TGGTCACTGCCCGCGAGACACAAACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 329
DB 1 TGGTCACTGCCCGCGAGACACACACCCAGGAGCATTTGAACACTGCACAGCGGCATCT 60

QY 330 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 367
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTAGCTA 98

RESULT 10
LOCUS A1529939 799 bp mRNA linear EST 18-MAR-1999
DEFINITION u167c09.y1 Sugano mouse liver mlia Mus musculus cDNA clone
IMAGE:1889392 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1529939 GI:4444074
 VERSION A1529939.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 799)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:189392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME18-FL3; Site 1: DraIII (CACTGCTG); Site 2: DraIII (CACCAGTGT); 1st strand cDNA was primed with an oligo(dT) primer [ATGCGCCCTTTTCTTTTCTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTCTACTGG], digested and cloned into distinct DraIII sites of the pME18-FL3 vector (5' site CACTGCTG, 3' site CACCAGTGT). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCACTGACGACACA."

ORIGIN
 Query Match 23.3%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 9.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 270 TGGTCACCTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGCGCCATCT 329
 |||||
 Db 1 TGGTCACCTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGCGCGTCT 60
 |||||
 Oy 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
 |||||
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTA 98
 |||||

RESULT 11
 CA478518 811 bp mRNA linear EST 09-MAR-2004
 LOCUS AGENCOURT 10789306 NIH_MGC_152 Mus musculus cDNA clone
 DEFINITION IMAGE:6766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 811)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LKMO0138 row: m column: 23
 High quality sequence stop: 536.
 location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCCTGACGCGCCGACACACCACTTTGACAGAAAGCTGGTCTTTTCTTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen. Library amplification 16 cycles. Library constructed by Mark Bittinger in the Bradfield Laboratory (McArdle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 23.3%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 9.8e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 270 TGGTCACCTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGCGCCATCT 329
 |||||
 Db 39 TGGTCACCTGCCCCCGAGACACAACCAGGAGCATTTGAACACTGACACGCGCGTCT 98
 |||||
 Oy 330 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTAGCTA 367
 |||||
 Db 99 GCCCAGAGAGCTGTGACCAACCACTTCGCTACTACTA 136
 |||||

RESULT 12
 A1790802 848 bp mRNA linear EST 02-JUL-1999
 LOCUS uk28b10.Y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 DEFINITION IMAGE:1970299 5', similar to gb:X8159.M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 848)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28b10.x1
 Contact: Maira M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:987039

FEATURES

source

Seq primer: custom primer used
 High quality sequence stop: 514.
 Location/Qualifiers

```
1..848
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970229"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse kidney mKia"
/notes="Organ: kidney; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCAGTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTTACTGTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTTAAAGCTGCG and 3' end
primer CGACCTGACGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 23.3%; Score 86.8; DB 1; Length 848;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 270 TGGTCCACATGCCCGCGGAGACACAAACCAGAGAGATGAACTGACACAGGCGCATCT 329
    |||
Db 1 TGGTCCACATGCCCGCGGAGACACAAACCAGAGAGATGAACTGACACAGGCGCATCT 60
    |||
QY 330 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 367
    |||
Db 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
    |||
```

RESULT 13

LOCUS

AI528304 852 bp mRNA linear EST 18-MAR-1999
 DEFINITION u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone

IMAGE:1890210.5, similar to gb:M59316.trai INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA
 sequence.

ACCESSION

VERSION

AI528304 AI528304.1 GI:4442439

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 852)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuk,R.,
 Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.

AUTHORS

The Washu-NCI Mouse EST Project 1999

TITLE

The Washu-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
 Contact: Maira M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:974534

FEATURES

source

Seq primer: custom primer used
 High quality sequence stop: 478.
 Location/Qualifiers

```
1..852
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mla"
/notes="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCAGTGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCTTACTGTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCAGTGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTTAAAGCTGCG and 3' end
primer CGACCTGACGCTCGAGCACA."
```

ORIGIN

Query Match

Best Local Similarity 23.3%; Score 86.8; DB 1; Length 852;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 270 TGGTCCACATGCCCGCGGAGACACAAACCAGAGAGATGAACTGACACAGGCGCATCT 329
    |||
Db 1 TGGTCCACATGCCCGCGGAGACACAAACCAGAGAGATGAACTGACACAGGCGCATCT 60
    |||
QY 330 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 367
    |||
Db 61 GCCCAGAGAGCTGTGACACCACTTCGCTACTAGCTA 98
    |||
```

RESULT 14

LOCUS

AI196154 605 bp mRNA linear EST 14-OCT-1998
 DEFINITION u195d08.y1 Sugano mouse liver mla Mus musculus cDNA clone

IMAGE:1887663.5, similar to gb:M59316.trai INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA
 sequence.

ACCESSION

VERSION

AI196154 AI196154.1 GI:3748760

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 605)
 Marra,M., Hillier,L., Kucaba,T., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.

AUTHORS

The Washu-HHMI Mouse EST Project

TITLE

The Washu-HHMI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:971987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

1..605

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTAAAGCTGG and 3' end
primer CGACCTGCAGCTCGACACA."

ORIGIN

Query Match

Best Local Similarity 22.9%; Score 85.2; DB 1; Length 605;
Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 270 TGTGTCACCTGCCCCGAGACACAACCCAGAGCATTTGAACATGCACACGGCCATCT 329
|||||
1 TGTGTCACCTGCCCCGAGACACAACCCAGAGCATTTGAACATGCACACGGCTCT 60

Db 1 TGTGTCACCTGCCCCGAGACACAACCCAGAGCATTTGAACATGCACACGGCTCT 60

Qy 330 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 367
|||||
61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98

Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98

RESULT 15
A1785039 380 bp mRNA linear EST 02-JUL-1999

LOCUS u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone

DEFINITION IMAGE:1888018.5, similar to gb:X81579.M.musculus mRNA for

insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1785039

VERSION A1785039.1 GI:5332755

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus (house mouse)

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pepe,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCam,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

1..380

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mlia"
/note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTAAAGCTGG and 3' end
primer CGACCTGCAGCTCGACACA."

ORIGIN

Query Match

Best Local Similarity 22.1%; Score 82.2; DB 1; Length 380;
Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 273 TCCACTGCCCCGAGACACAACCCAGAGCATTTGAACATGCACACGGCCATCTG 332
|||||
1 TCCACTGCCCCGAGAGACACAACCCAGAGCATTTGAACATGCATGCGCGCTCTCC 60

Db 1 TCCACTGCCCCGAGAGACACAACCCAGAGCATTTGAACATGCATGCGCGCTCTCC 60

Qy 333 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 367
|||||
61 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 95

Db 61 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 95

Search completed: September 1, 2005, 06:56:52
Job time : 2831.35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12; Search time 1083.43 Seconds
(without alignments)
2248.600 Million cell updates/sec

Title: US-09-972-916b-5
Perfect score: 372
Sequence: 1 tacactggggggccagatcc.....ttccgctactagtagcgagc 372

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	372	100.0	372	US-09-972-916A-5	Sequence 5, Appli
2	321	86.3	321	US-09-972-916A-4	Sequence 6, Appli
3	223.6	60.1	423	US-09-972-916A-6	Sequence 4, Appli
4	219.4	59.0	270	US-09-972-916A-3	Sequence 3, Appli
5	219	58.9	219	US-09-972-916A-2	Sequence 2, Appli
6	154	41.4	423	US-09-972-916A-6	Sequence 6, Appli
7	76.4	20.5	1500	US-09-917-800A-1608	Sequence 1608, Ap

8	76.4	20.5	1500	17	US-10-388-934-5	Sequence 5, Appli
9	76.4	20.5	1500	17	US-10-191-803-73	Sequence 73, Appl
10	76.4	20.5	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
11	63.8	17.2	6128	9	US-09-880-107-2393	Sequence 2393, Ap
12	63.8	17.2	6128	22	US-10-756-149-1484	Sequence 1484, Ap
13	63.8	17.2	9173	22	US-10-893-315-126	Sequence 126, App
14	63.8	17.2	9174	22	US-10-893-315-160	Sequence 160, App
15	55.2	14.8	13011	17	US-10-388-934-36	Sequence 36, Appli
16	55.2	14.8	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
17	51	13.7	51	9	US-09-972-916A-1	Sequence 1, Appli
18	51	13.7	270	9	US-09-972-916A-3	Sequence 3, Appli
19	40.2	10.8	31703	17	US-10-085-117-172	Sequence 172, App
20	37.8	10.2	7061	15	US-10-311-455-970	Sequence 970, App
21	37.8	10.2	7061	17	US-10-221-613-148	Sequence 148, App
22	37.6	10.1	3191	16	US-10-029-386-20306	Sequence 20306, A
23	37.4	10.1	12850	14	US-10-017-122-1	Sequence 1, Appli
24	36.8	9.9	761	17	US-10-264-237-828	Sequence 828, App
25	36.6	9.8	1678	17	US-10-094-749-9	Sequence 9, Appli
26	36.2	9.7	586	16	US-10-029-386-5186	Sequence 5186, Ap
27	35.8	9.6	11009	9	US-09-845-583-1	Sequence 1, Appli
28	35.8	9.6	11009	14	US-10-037-182-3	Sequence 1, Appli
29	35.8	9.6	11009	21	US-10-764-420-947	Sequence 947, App
30	35.6	9.6	490	10	US-09-918-995-22070	Sequence 22070, A
31	35.4	9.5	1934	18	US-10-424-599-61023	Sequence 61023, A
32	35.2	9.5	1687	20	US-10-739-930-5461	Sequence 5461, Ap
33	35	9.4	769	13	US-10-027-632-164336	Sequence 164336, A
34	35	9.4	769	17	US-10-027-632-164336	Sequence 164336, A
35	35	9.4	3895	14	US-10-011-585A-76	Sequence 76, Appli
36	34.4	9.2	985	19	US-10-767-701-15088	Sequence 15088, A
37	34	9.1	10537	9	US-09-764-869-1269	Sequence 1269, Ap
38	34	9.1	10537	17	US-10-091-504-1269	Sequence 1269, Ap
39	34	9.1	10537	17	US-10-227-577-1269	Sequence 1270, Ap
40	34	9.1	10543	9	US-09-764-869-1270	Sequence 1270, Ap
41	34	9.1	10543	14	US-10-091-504-1270	Sequence 1270, Ap
42	34	9.1	10543	17	US-10-227-577-1270	Sequence 1270, Ap
43	34	9.1	49979	19	US-10-741-601-5746	Sequence 5746, Ap
44	34	9.1	49979	21	US-10-741-600-17905	Sequence 17905, A
45	33.4	9.0	201	20	US-10-719-993-35505	Sequence 35505, A

ALIGNMENTS

RESULT 1
US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 100.0%; Score 372; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 2, 5e-116;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACACTGGGGGCGAGATCCAGAAACACGAGAGTCCCGCCCATTTACACTGGG 60
|||
DB 1 TACACTGGGGGCGAGATCCAGAAACACGAGAGTCCCGCCCATTTACACTGGG 60
61 GGCAGAGTCCAGAACACGAGAGTCCCGCCCATTTACACTGGGGGCGAGAGT 120

```

Db      61 GGCAGAGTCCAGAAACAAGGAGTGCCTGGCCCATGATACCTGGGGCCAGAGT 120
Qy      121 CAGAGAACCAAGGGAGTGCCTGGCCCATGATGACAAACAACTATTGTA 180
Db      121 CAGAGAACCAAGGGAGTGCCTGGCCCATGATGACAAACAACTATTGTA 180
Qy      181 CAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGGAGCAGCCCTT 240
Db      181 CAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGGAGCAGCCCTT 240
Qy      241 CATAAAGCCCTGGATATGCGCAAGCAGATGATGCACTGCGCGAGACAAACCCAG 300
Db      241 CATAAAGCCCTGGATATGCGCAAGCAGATGATGCACTGCGCGAGACAAACCCAG 300
Qy      301 CAGACATTGAACATCTGACACAGGCGCATCTGCGCCAGAGAGCTGTGACCACTTCGCTA 360
Db      301 CAGACATTGAACATCTGACACAGGCGCATCTGCGCCAGAGAGCTGTGACCACTTCGCTA 360
Qy      361 CTAGCTAGCCGC 372
Db      361 CTAGCTAGCCGC 372

```

```

RESULT 2
US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIORITY FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIORITY FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-4

```

```

Query Match      86.3%; Score 321; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 5.8e-99;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      52 TACACTGGGGGCGAAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACCTGGG 111
Db      1 TACACTGGGGGCGAAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACCTGGG 60
Qy      112 GGCAGAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACAAACAACT 171
Db      61 GGCAGAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACAAACAACT 120
Qy      172 TATTTTGAACAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGGAG 231
Db      121 TATTTTGAACAGGGGATCTAGACGCTGCGCTGACAAATTAACCCGCTGCGGAG 180
Qy      232 CAGGCTTCAATTAAGGCTTGGATATGCGCAAGCAGATGATGCACTGCGCGAGACA 291
Db      181 CAGGCTTCAATTAAGGCTTGGATATGCGCAAGCAGATGATGCACTGCGCGAGACA 240
Qy      292 CAACCCAGCAGATTAACAATGACACAGGCGCATCTGCGCCAGAGAGCTGTGACCA 351
Db      241 CAACCCAGCAGATTAACAATGACACAGGCGCATCTGCGCCAGAGAGCTGTGACCA 300
Qy      352 CTTCGCTACTAGCTAGCCGC 372
Db      301 CTTCGCTACTAGCTAGCCGC 321

```

```

RESULT 3
US-09-972-916A-6
; Sequence 6, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIORITY FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIORITY FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 6
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-6

```

```

Query Match      60.1%; Score 223.6; DB 9; Length 423;
Best Local Similarity 76.2%; Pred. No. 9.6e-66;
Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

```

```

Qy      2 AACTGGGGGCGAAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACCTGGG 61
Db      34 ACTGTGCCCCCACTGATGACATGAGGCGGACAGGAGCATCCGCTGATCTGATCTGACC 93
Qy      62 GCCAGAGTCCAGAAACAAGGAGTGCCTGGCCCATGATGACCTGGGGCCAGAGTGC 121
Db      94 CCGAGTGAACATGAGGCGGACAGGAGCATCCCGTGTGCTGTGATCTGCCCCAGAGTGA 153
Qy      122 CAGAACCAAGGAGTGCCTGGCCCATG-----TCAAGCA 162
Db      154 CATGGGCGACAGGGGCACTCCGCTGTCTGACTGTGCCCCCACTGATGACAAAGCA 213
Qy      163 AAACAACTATTTTGAACAGGGGATCTAGACGCTGCGCCATGACATTAACCGCT 222
Db      214 AAACAACTATTTTGAACAGGGGATCTAGACGCTGCGCCATGACATTAACCGCT 273
Qy      223 GCTCCGAGCAGCCCTTCAATTAAGGCTTGGATATGCGCAAGCAGATGATGCACTGCC 282
Db      274 GCTCCGAGCAGCCCTTCAATTAAGGCTTGGATATGCGCAAGCAGATGATGCACTGCC 333
Qy      283 GCCGAGACAAACCCAGCAGATTAACAATGACACAGGCGCATCTGCGCCAGAGAGCTG 342
Db      334 GCCGAGACAAACCCAGCAGATTAACAATGACACAGGCGCATCTGCGCCAGAGAGCTG 393
Qy      343 TGACCACTACTTCGCTACTAGCTAGCCGC 372
Db      394 TGACCACTACTTCGCTACTAGCTAGCCGC 423

```

```

RESULT 4
US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972, 916A
; PRIORITY FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239, 113
; PRIORITY FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Synthesized
US-09-972-916A-3

Query Match 59.0%; Score 219.4; DB 9; Length 270;
Best Local Similarity 99.5%; Pred. No. 2.3e-64;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 152 TGTCAAGCAAAACAACCTATTGTAACAGGGGATCTGACAGCGCTGACCAT 211
DB 50 TATCAAGCAAAACAACCTATTGTAACAGGGGATCTGACAGCGCTGACCAT 109
QY 212 CATTAACCCGCTGCTGCGAGCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCATG 271
DB 110 CATTAACCCGCTGCTGCGAGCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCATG 169
QY 272 GTTCACTGCGCGCGAGACACAACCAGCAGCATTTGAACACTGACACGGCCATCTGC 331
DB 170 GTTCACTGCGCGCGAGACACAACCAGCAGCATTTGAACACTGACACGGCCATCTGC 229
QY 332 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGC 372
DB 230 CCAGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGC 270

RESULT 5
US-09-972-916A-2

Sequence 2, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 2

LENGTH: 219

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

US-09-972-916A-2

Query Match 58.9%; Score 219; DB 9; Length 219;
Best Local Similarity 100.0%; Pred. No. 3e-64;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 TCACAGCAAAACAACCTATTGTAACAGGGGATCTGACAGCGCTGACCATCA 213
DB 1 TCACAGCAAAACAACCTATTGTAACAGGGGATCTGACAGCGCTGACCATCA 60
QY 214 TTAACCCGCTGCTGCGAGCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCATG 273
DB 61 TTAACCCGCTGCTGCGAGCAGCCCTTCATTAAGCCCTGGGTATGCGCAGCATG 120
QY 274 CCAGTCCCGCGCGAGACACAACCAGCAGCATTTGAACACTGACACGGCCATCTGCC 333
DB 121 CCAGTCCCGCGCGAGACACAACCAGCAGCATTTGAACACTGACACGGCCATCTGCC 180
QY 334 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGC 372
DB 181 AGAGAGCTGTGACCAACCACTTCGCTACTAGTACCGC 219

RESULT 6
US-09-972-916A-6/c

Sequence 6, Application US/09972916A

Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

CURRENT FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 6

LENGTH: 423

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized

US-09-972-916A-6

Query Match 41.4%; Score 154; DB 9; Length 423;
Best Local Similarity 100.0%; Pred. No. 5e-42;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACACTGGGGGCGAGATTCAGAAACACGGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 60
DB 204 TACACTGGGGGCGAGATTCAGAAACACGGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 145
QY 61 GCGCAGATCCAGGAACACGGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 120
DB 144 GCGCAGATCCAGGAACACGGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 85
QY 121 CCAGAACCAAGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 154
DB 84 CCAGAACCAAGGAGTCCCGTGCCTATGACTGGGGCCAGAGT 51

RESULT 7

US-09-917-800A-1608

Sequence 1608, Application US/09917800A

Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Caetle, Arthur

APPLICANT: Elshoff, Michael

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5018-US

CURRENT APPLICATION NUMBER: US/09/917,800A

PRIOR FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303,459

PRIOR FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1608

LENGTH: 1500

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_013144

US-09-917-800A-1608

PRIOR APPLICATION NUMBER: US 60/377,611
PRIOR FILING DATE: 2002-05-06

US-10-152-319A-1

US-10-152-319A-1613	20.5%	Score 76.4;	DB 18;	length 1500;
Query Match	20.5%	Score 76.4;	DB 18;	length 1500;
Query Match	20.5%	Score 76.4;	DB 18;	length 1500;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 117.486 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916b-5

Perfect score: 372
Sequence: 1 taccatggggggccagagcgc.....ttccgctactagctagcgcgc 372

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/pdata/1/ina/5A_COMB.seq:*
2: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
3: /cgn2_6/pdata/1/ina/5A_COMB.seq:*
4: /cgn2_6/pdata/1/ina/5B_COMB.seq:*
5: /cgn2_6/pdata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/pdata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63.8	17.2	9173	4 US-09-949-001-30	Sequence 30, Appl
2	63.8	17.2	9174	4 US-09-949-001-36	Sequence 36, Appl
3	55.2	14.8	194	4 US-08-945-140-1	Sequence 1, Appl
4	55.2	14.8	13011	2 US-08-791-849A-14	Sequence 14, Appl
5	40.2	10.8	15108	4 US-09-949-016-11786	Sequence 11786, A
6	40.2	10.8	15108	4 US-09-949-016-117205	Sequence 11705, A
7	36.6	9.8	3054	4 US-09-484-970B-138	Sequence 138, App
8	36.2	9.7	35471	4 US-09-949-016-12957	Sequence 12957, A
9	36	9.7	40586	4 US-09-949-016-16965	Sequence 16965, A
10	35.8	9.6	11009	4 US-09-845-583A-1	Sequence 1, Appl
11	35.8	9.6	16520	4 US-09-949-016-14710	Sequence 14710, A
12	35.8	9.6	16520	4 US-09-949-016-15394	Sequence 15394, A
13	35.6	9.6	76767	4 US-09-949-016-12147	Sequence 12147, A
14	35.6	9.6	76767	4 US-09-949-016-17361	Sequence 17361, A
15	35	9.4	601	4 US-09-949-016-19194	Sequence 19194, A
16	35	9.4	601	4 US-09-949-016-73784	Sequence 73784, A
17	35	9.4	2293	4 US-09-949-016-2153	Sequence 2153, Ap
18	35	9.4	2408	4 US-09-949-016-74	Sequence 74, Appl
19	35	9.4	7561	4 US-09-949-016-11816	Sequence 11816, A
20	35	9.4	7562	4 US-09-949-016-13895	Sequence 13895, A
21	34	9.1	30244	4 US-09-949-016-13208	Sequence 13208, A
22	34	9.1	30245	4 US-09-949-016-13550	Sequence 13550, A
23	33.4	9.0	364	4 US-09-270-767-28006	Sequence 28006, A
24	33.4	9.0	601	4 US-09-949-016-35649	Sequence 35649, A
25	33.4	9.0	601	4 US-09-949-016-177807	Sequence 177807, A
26	33.4	9.0	761	4 US-09-270-767-13109	Sequence 13109, A
27	32.8	8.8	749	3 US-09-257-583-12	Sequence 12, Appl

C 28	32.8	8.8	2130	4 US-09-909-962A-6	Sequence 6, Appl
C 29	32.8	8.8	2130	4 US-09-909-962A-7	Sequence 7, Appl
C 30	32.6	8.8	907	4 US-09-585-645A-55	Sequence 65, Appl
C 31	32.6	8.8	134008	4 US-09-949-016-13841	Sequence 13841, A
C 32	32.2	8.7	505	4 US-09-621-976-15639	Sequence 15639, A
C 33	32.2	8.7	601	4 US-09-949-016-35650	Sequence 35650, A
C 34	32.2	8.7	601	4 US-09-949-016-35651	Sequence 35651, A
C 35	32.2	8.7	601	4 US-09-949-016-177808	Sequence 177808, A
C 36	32.2	8.7	601	4 US-09-949-016-177809	Sequence 177809, A
C 37	32.2	8.7	1281	4 US-09-902-540-3669	Sequence 3669, Ap
C 38	32.2	8.7	4447	2 US-08-304-309-3	Sequence 3, Appl
C 39	32.2	8.7	4447	3 US-08-991-942-3	Sequence 3, Appl
C 40	32.2	8.7	23847	4 US-09-902-540-1177	Sequence 1177, Ap
C 41	32.2	8.7	46343	4 US-09-949-016-16824	Sequence 16824, A
C 42	32	8.6	36180	4 US-09-949-016-11745	Sequence 11745, A
C 43	32	8.6	36181	4 US-09-949-016-16163	Sequence 16163, A
C 44	31.8	8.5	2194	4 US-09-270-767-10411	Sequence 10411, A
C 45	31.6	8.5	1348	4 US-09-774-528-404	Sequence 404, App

ALIGNMENTS

```
RESULT 1
US-09-949-001-30
; Sequence 30, Application US/09949001
; Patent No. 6623336
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00789
; CURRENT APPLICATION NUMBER: US/09/949,001
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FASTSQ for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-30

Query Match      17.2%; Score 63.8; DB 4; Length 9173;
Best Local Similarity 63.2%; Pred. No. 2.2e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCACGAAACCAAGGAGTCCCGGCGCCATGTCACAAACAAACTTATTGGA 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1842 TCCTCCACCAAGCGGTTGCTGAGGCGCTTGAGTGCATTACCAAACTTATTGGA 1901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 180 ACACGGGAGCTCTAGACAGCTGCTGCAATCAATTAACCC-----GTGCTGCGG 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1902 ACACCTCAGCTCTGAGCGTGGCGGCGCTGCAATCAATCACTTCGTGCAAGTGGCGGG 1961
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 AGCAGCCTTCATPAGGCGCTGAGTATGCGCAGCAGCATGCTTCACTGCGCGGAGA 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1962 CCTGCGCCTTATPAGTGGCGCGCTGTGTCCACGAGCATGGCCACCGCCATCC---- 2017
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 CACAAACCCAGCAGCATGTAACCTGC-ACACGGCCATTCGCCAGAGACTGTGACCA 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2018 ----CATCAGCAGCATCTGCGCGCGCGCCGACCCCTCCAGAGACTGTGACCA 2073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 349 CCACTTC 355
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2074 CCGCTCC 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-949-001-36
; Sequence 36, Application US/09949001
; Patent No. 6625336
```

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL000789
CURRENT APPLICATION NUMBER: US/09/949,001
CURRENT FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: 60/231,323
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 848
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 36
LENGTH: 9174
TYPE: DNA
ORGANISM: Human
US-09-949-001-36

Query Match 17.2%; Score 63.8; DB 4; Length 9174;
Best Local Similarity 63.2%; Pred. No. 2.2e-10;
Matches 156; Conservative 0; Mismatches 72; Indels 19; Gaps 3;

QY 120 TCCAGAACACGGAGTGCCTCCGCGCCATGTCACAGCAAAACAATTATTGGA 179
DB 1842 TCCCTCCACACAGCGGTTGCGGTAGGGGCTTGGGTGCACATGCAAAACAATTATTGGA 1901
QY 180 ACACGGGATCTAGACAGCGCTGCCCTGACATCATTAACCC-----GTGCTGCGG 229
DB 1902 ACACCTCAGCTCTAGACAGCGCGCGCTGCCATCATTAACCTCTGTCAGATGCGCGG 1961
QY 230 AGCAGCCCTTCATTAAGCGCTGCGGTATGCGCAGCAGCATGTCACCTGCCGCGGAGA 289
DB 1962 CCGTGGCCCTTATTAAGTGGCGCGCTGTGTCCAGGAGATGCGCACCGCCATC----- 2017
QY 290 CACAACCCAGGACATTTGAACACTGC-ACACGCCCATCTGCCAGAGAGTGTGACCA 348
DB 2018 ---CATCAGGAGCATCTGCGCGCGCGCGCCACCCCTCCAGAGAGCATGCGCA 2073
QY 349 CCACTTC 355
DB 2074 CCGCTCC 2080

RESULT 3
US-08-945-140-1/C
Sequence 1, Application US/08945140
Patent No. 6309878
GENERAL INFORMATION:
APPLICANT: CHEN, Ruihuan
APPLICANT: DOIRON, Bruno
APPLICANT: KAHN, Axel
TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
TITLE OF INVENTION: VECTOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,140
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 95/04558
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR96/00560
FILING DATE: 12-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky Beg., Martin F.
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: EX95002-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 194 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-945-140-1

Query Match 14.8%; Score 55.2; DB 3; Length 194;
Best Local Similarity 88.2%; Pred. No. 2.6e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 35 TGCCCCGTGCGCCCATGTACACTGCGGCGCAGAGTCCAGAACCGGAGTCCCCGTG 94
DB 78 TGCCACGGAAGCCTGTGACACTGCGGCGCAGAGTCCAGAACCGGAGTCCCCGTG 19
QY 95 CGCCCATG 102
DB 18 CGCCCATG 11

RESULT 4
US-08-791-849A-14/C
Sequence 14, Application US/08791849A
Patent No. 591449
GENERAL INFORMATION:
APPLICANT: Makoto MURASE et al.
TITLE OF INVENTION: Method for Increasing Storage
TITLE OF INVENTION: Lipid Content in Plant Seed
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,849A
FILING DATE: January 30, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 13011 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

```

MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: rat (Rattus norvegicus)
FEATURE:
NAME/KEY: CDS
LOCATION: join(3212..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077
FEATURE:
NAME/KEY: intron
LOCATION: join(3219..3765, 3949..5916, 6009..6151,
6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297
LOCATION: 9480..10162)
US-08-791-849A-14

```

```

Query Match 14.8% Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.7e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 35 GTGCCCCGTGCGCCATGTACCTGGGGGCGAGTCCAGAACCGAGGAGTCCCGGTG 94
DB 3088 TGCACAGGAGCCTGTATGACTGGGGGCGAGTCCAGAACCGAGGAGTCCCGGTG 3029
QY 95 CGCCCATG 102
DB 3028 CGCCCATG 3021

```

```

RESULT 5
US-09-949-016-11786/c
Sequence 11786, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11786
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-11786

```

```

Query Match 10.8% Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 34 GTGCCCCGTGCGCCATGTACCTGGGGGCGAGTCCAGAACCGAGGAGTCCCGGTG 93
DB 2417 GTTCCATTGGAAGCCCTGTATGCGAGGGGCGAGTCCAGAACCGAGGAGTCCCGGT 2358
QY 94 G 94
DB 2357 G 2357

```

RESULT 6
US-09-949-016-17205/c

```

Sequence 17205, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17205
LENGTH: 15108
TYPE: DNA
ORGANISM: Human
US-09-949-016-17205

```

```

Query Match 10.8% Score 40.2; DB 4; Length 15108;
Best Local Similarity 78.7%; Pred. No. 0.015;
Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 34 GTGCCCCGTGCGCCATGTACCTGGGGGCGAGTCCAGAACCGAGGAGTCCCGGTG 93
DB 2417 GTTCCATTGGAAGCCCTGTATGCGAGGGGCGAGTCCAGAACCGAGGAGTCCCGGT 2358
QY 94 G 94
DB 2357 G 2357

```

```

RESULT 7
US-09-484-970B-138/c
Sequence 138, Application US/09484970B
Patent No. 6426186
GENERAL INFORMATION:
APPLICANT: Jones, Karen A.
APPLICANT: Volkmutch, Wayne
APPLICANT: Walker, Michael G.
TITLE OF INVENTION: BONE REMODELING GENES
FILE REFERENCE: PB-0014 US
CURRENT APPLICATION NUMBER: US/09/484,970B
CURRENT FILING DATE: 2000-01-18
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PERL Program
SEQ ID NO 138
LENGTH: 3054
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. 6426186 336987.1CB1
US-09-484-970B-138

```

```

Query Match 9.8% Score 36.6; DB 3; Length 3054;
Best Local Similarity 62.6%; Pred. No. 0.11;
Matches 57; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 2 ACACGTGGGGGCGAGTCCAGAACCGAGGAGTCCCGGTGCGCCATGTACCTGGGGG 61
DB 355 ACTTCGGAAGGAGAGCCCTGTATGCGAGGGGCGAGTCCAGAACCGAGGAGTCCCGGT 296
QY 62 GCCAGATCCAGAACCGAGGAGTCCCGGTG 92
DB 295 GCTGAGGCTTGGAGCGAAGGAGAGGCCCGG 265

```

RESULT 8

US-09-949-016-12957/c
; Sequence 12957, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12957
; LENGTH: 35471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12957

Query Match 9.7%; Score 36.2; DB 4; Length 35471;
Best Local Similarity 57.5%; Pred. No. 0.45;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 22 GGAACCAAGGAGTGGCCCGGCGCCATGTACACTGGGGGCGAGAGTCCAGAACACAG 81
Db 22896 GGCAGACAGGAGTCTCCCGGAGGACACAGCTCCCGGCGCCAGGAGGAGGAGGCGG 22837
Qy 82 GAGTGGCCCGTGGCCCATGTACACTGGGGGCGAGAGTCCAGAACACAGG 134
Db 22836 GCAGGCGGAGTCCCGGAGGACACAGCTCCCGGAGGAGGAGGAGGAGG 22784

RESULT 9
US-09-949-016-16965
; Sequence 16965, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16965
; LENGTH: 40586
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16965

Query Match 9.7%; Score 36; DB 4; Length 40586;
Best Local Similarity 52.7%; Pred. No. 0.56;
Matches 78; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 2 ACACATGGGGGCGAGTCCAGAACACGAGGAGTGGCCCGGCGCCATGTACACTGGGG 61
Db 25936 ACTGTGGGAGACCTCGCTGACCCACTCTCCCGGAGGAGGAGGAGGAGGAGG 25995

Qy 62 GCACAGTCCAGGAACACGAGGAGTGGCCCGGCGCCATGTACACTGGGGGCGAGAGTCC 121
Db 25996 ACACCTACTGTGACCCACCTCCCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 26055
Qy 122 CAGGAACCAAGGAGTGGCCCGGCGCC 149
Db 26056 CTGACCCACCTCTCCCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 26083

RESULT 10
US-09-845-583A-1
; Sequence 1, Application US/09845583A
; Patent No. 6635616
; GENERAL INFORMATION:
; APPLICANT: Burgess, Robert
; APPLICANT: Brunken, William Joseph
; APPLICANT: Champilaud, Marie-France
; APPLICANT: Hunter, Dale
; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
; FILE REFERENCE: 10287-056001
; CURRENT APPLICATION NUMBER: US/09/845,583A
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/200,863
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 11009
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-845-583A-1

Query Match 9.6%; Score 35.8; DB 4; Length 11009;
Best Local Similarity 53.1%; Pred. No. 0.36;
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy 3 CACTGGGGGCGAGTCCAGGAACACGAGGAGTGGCCCGGCGCCATGTACACTGGGGG 62
Db 6022 CACTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6081
Qy 63 CCAAGTCCAGGAACACGAGGAGTGGCCCGGCGCCATGTACACTGGGGGCGAGAGTCC 122
Db 6082 CTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6141
Qy 123 AGGAACCAAGGAGTGGCCCGGCG 145
Db 6142 GGCACCTGACCTGTCTCCCGGAG 6164

RESULT 11
US-09-949-016-14710
; Sequence 14710, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14710
; LENGTH: 16520
; TYPE: DNA
; ORGANISM: Human

US-09-949-016-14710

Query Match 9.6%; Score 35.8; DB 4; Length 16520;
 Best Local Similarity 52.3%; Pred. No. 0.43;
 Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 7 GGGGGCCAGAGTCCAGGAACCAAGAGTGCCTGCGCCCATGTACACTGGGGCCAG 66
 DB 3583 GGGTGTGGGTGTCCCGAGAGTGTGTCCTCGGGGCGTGGTGTCCCGAGAGTGGG 3642
 QY 67 AGTCCAGAACCAACGAGAGTGCCTGCGCCCATGTACACTGGGGCCAGAGTCCAGGA 126
 DB 3643 TGTCCCGGGGCGGTGGTGTCCCGAGAGTGTGAGTGTCCCGGGGCGTGGTGTCCCGGG 3702
 QY 127 ACCACGGGAGTGCCTGCGCCCATGTAC 157
 DB 3703 AGTGTGGTGTCCCGGGGCGTGGTGTCC 3733

RESULT 12

US-09-949-016-15394
 ; Sequence 15394, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15394
 ; LENGTH: 16520
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-15394

Query Match 9.6%; Score 35.8; DB 4; Length 16520;
 Best Local Similarity 52.3%; Pred. No. 0.43;
 Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 7 GGGGGCCAGAGTCCAGGAACCAAGAGTGCCTGCGCCCATGTACACTGGGGCCAG 66
 DB 3583 GGGTGTGGGTGTCCCGAGAGTGTGTCCTCGGGGCGTGGTGTCCCGAGAGTGGG 3642
 QY 67 AGTCCAGAACCAACGAGAGTGCCTGCGCCCATGTACACTGGGGCCAGAGTCCAGGA 126
 DB 3643 TGTCCCGGGGCGGTGGTGTCCCGAGAGTGTGAGTGTCCCGGGGCGTGGTGTCCCGGG 3702
 QY 127 ACCACGGGAGTGCCTGCGCCCATGTAC 157
 DB 3703 AGTGTGGTGTCCCGGGGCGTGGTGTCC 3733

RESULT 13

US-09-949-016-12147/c
 ; Sequence 12147, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755

;; PRIOR FILING DATE: 2000-10-20
 ;; PRIOR APPLICATION NUMBER: 60/237,768
 ;; PRIOR FILING DATE: 2000-10-03
 ;; PRIOR APPLICATION NUMBER: 60/231,498
 ;; PRIOR FILING DATE: 2000-09-08
 ;; NUMBER OF SEQ ID NOS: 207012
 ;; SOFTWARE: FastSeq for Windows Version 4.0
 ;; SEQ ID NO 12147
 ;; LENGTH: 767677
 ;; TYPE: DNA
 ;; ORGANISM: Human
 ;; FEATURE:
 ;; NAME/KEY: misc.feature
 ;; LOCATION: (1)...(767677)
 ;; OTHER INFORMATION: n = A,T,C or G
 ; US-09-949-016-12147

Query Match 9.6%; Score 35.6; DB 4; Length 767677;
 Best Local Similarity 51.9%; Pred. No. 2.8; Indels 0; Gaps 0;
 Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 7 GGGGGCCAGAGTCCAGGAACCAAGAGTGCCTGCGCCCATGTACACTGGGGCCAG 66
 DB 562461 GGGGCGCTGAGCAAGAGATCAAGGACTCCCGCACCCAGTCCACAGGAGCCCTG 562402
 QY 67 AGTCCAGAACCAACGAGAGTGCCTGCGCCCATGTACACTGGGGCCAGAGTCCAGGA 126
 DB 562401 AGTGTGAGATCAACGAGACTCCTTACACCTTACACTGACAGGGGCGCTTGAGCAAGAG 562342
 QY 127 ACCACGGGAGTGCCTGCGCCCATGTAC 160
 DB 562341 ATCAGGAGACTCCTTGCACCCACAGTCCACG 562308

RESULT 14

US-09-949-016-17361/c
 ; Sequence 17361, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17361
 ; LENGTH: 767677
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(767677)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-949-016-17361

Query Match 9.6%; Score 35.6; DB 4; Length 767677;
 Best Local Similarity 51.9%; Pred. No. 2.8;
 Matches 80; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 7 GGGGGCCAGAGTCCAGGAACCAAGAGTGCCTGCGCCCATGTACACTGGGGCCAG 66
 DB 562461 GGGGCGCTGAGCAAGAGATCAAGGACTCCCGCACCCAGTCCACAGGAGCCCTG 562402
 QY 67 AGTCCAGAACCAACGAGAGTGCCTGCGCCCATGTACACTGGGGCCAGAGTCCAGGA 126

Db 562401 AGTGTGAGATACGGGACTCCCTACACCTACAGTACTACGGGGCCTGAGCAAGAG 562342
 QY 127 ACCACGGAGTGCCCGTGGCCCATGTCAAG 160
 Db 562341 ATCACGGGACTCCCTGACACCCACAGTCCACAG 562308

RESULT 15
 US-09-949-016-19194/C
 ; Sequence 19194, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CI001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 19194
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-19194

Query Match 9.4%; Score 35; DB 4; Length 601;
 Best Local Similarity 55.3%; Pred. No. 0.18;
 Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
 QY 232 CCAGCCCTTCATTAAGGCCCTGGGTATGGCCAGCCAGCATGTCTCACTGCCCGCCGAGACA 291
 Db 478 CCAGGCTTCTCGGGCGCTGTTCCTCTCATCCAGCTTAGCCACCTGCGAGTTAGGCC 419
 QY 292 CAAACCCAGCGGACATTGAACCTGCACAGGGCATCTGCCAGAGAGCTGTGACCACCA 351
 Db 418 CCCATCCAGAGACACATCCACCGGATCACAGCCCTGCCACACCACTTCATGATAGCC 359
 QY 352 CTT 354
 Db 358 CTT 356

Search completed: September 1, 2005, 07:07:16
 Job time : 123.652 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: August 31, 2005, 13:10:38 ; Search time 2633.79 Seconds
(without alignments)
7782.159 Million cell updates/sec

Title: US-09-972-916b-6

Perfect score: 423
Sequence: 1 celtsgggcgacggggcgacac.....ttccgctactagctagccgc 423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hgt: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_srs: *
12: gb_sy: *
13: gb_un: *
14: gb_vt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	212.4	50.2	1181	10	RAT1IGFZ M84484 Rattus norv
2	212.4	50.2	185148	2	AC136382 Rattus no
3	188.4	44.5	1363	10	MM1IGF X67493 M. musculus
4	188.4	44.5	192843	10	AL607124 Mouse DNA
5	98.2	23.2	1569	10	BC013345 Mus muscu
6	76.6	18.1	1507	10	AYS60836 Rattus norv
7	76.4	18.1	1500	6	AX401932 Rattus norv
8	76.4	18.1	1500	6	AX827271 Rattus norv
9	76.4	18.1	1500	10	RAT1IGF M84484 Rattus norv
10	76.4	18.1	1500	6	AX163782 Rattus norv
11	76.4	18.1	1500	10	RAT1IGF M84484 Rattus norv
12	65.4	15.5	1510	10	BC078889 Rattus norv
13	62.6	14.8	3886	9	AY095345 Papio anu
14	60.8	14.4	448	11	G67139 G67139 IGFBI X1.1
15	60.8	14.4	6128	6	AX409747 Rattus norv
16	60.8	14.4	6128	9	HUMIGFPIA M74587 Human insul
17	60.8	14.4	6128	11	G19994 G19994 sWSS752 Eri
18	60.8	14.4	6480	9	HUMIGFPIA M59316 Human insul
19	60.8	14.4	9082	9	AY434089 Homo sapi

20	60.8	14.4	69887	9	AC091524 AC091524 Homo sapi
21	59.2	14.0	141539	2	AC146152 AC146152 Pan trogl
22	59.2	14.0	189932	2	AC146117 AC146117 Pan trogl
23	59.2	14.0	200935	2	AC148834 AC148834 Pan trogl
24	55.8	13.2	125020	2	AF429315 AF429315 Homo sapi
25	55.8	13.2	256781	2	AC097952 AC097952 Rattus no
26	55.8	13.2	259329	2	AC109570 AC109570 Rattus no
27	55.2	13.0	194	6	A57715 A57715 Sequence 1
28	55.2	13.0	194	6	AR175909 AR175909 Sequence
29	55.2	13.0	13011	6	E14395 E14395 gDNA encod1
30	55.2	13.0	13011	6	AX827302 AX827302 Sequence
31	55.2	13.0	13011	10	RMLPKG X05684 Rat L-PK ge
32	53.6	12.7	81704	9	AC110299 AC110299 Homo sapi
33	53.6	12.7	151700	9	AC133528 AC133528 Homo sapi
34	53.4	12.6	200340	9	AC100797 AC100797 Homo sapi
35	52.2	12.3	170654	10	AC116502 AC116502 Mus muscu
36	52.2	12.3	187397	2	AC114566 AC114566 Mus muscu
37	52.2	12.3	218657	10	AC110517 AC110517 Mus muscu
38	51.6	12.2	184889	9	AL442125 AL442125 Human DNA
39	50.6	12.0	994	11	BY006814 BY006814 STS_C8B24
40	50.6	12.0	107103	9	AL589702 AL589702 Human DNA
41	50.6	12.0	145540	9	AP001052 AP001052 Homo sapi
42	50.6	12.0	184104	9	AC020709 AC020709 Homo sapi
43	50.6	12.0	340000	9	AP001752 AP001752 Homo sapi
44	50.4	11.9	895	9	HUMIGFPIA M23592 Human insul
45	50	11.8	53370	9	AL592071 AL592071 Human DNA

ALIGNMENTS

RESULT 1	RAT1IGFZ	1181 bp	DNA	linear	ROD 27-APR-1993
LOCUS	Rattus norvegicus insulin-like growth factor gene fragment.				
DEFINITION	M84484.1 GI:204927				
ACCESSION	insulin-like growth factor.				
VERSION	Rattus norvegicus (Norway rat)				
KEYWORDS	Rattus norvegicus				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;				
REFERENCE	Rattus.				
AUTHORS	1 (bases 1 to 1181)				
JOURNAL	Uncerman, T.G., Jackson, R.G., McGary, E., Whalen, C. and Goswami, R.G.				
COMMENT	Biochem. Biophys. Res. Commun. (1991) In press				
FEATURES	Original source text: Rattus norvegicus (strain Sprague-Dawley)				
source	male adult liver DNA.				
location/Qualifiers					
1..1181					
/organism="Rattus norvegicus"					
/mol_type="genomic DNA"					
/strain="Sprague-Dawley"					
/db_xref="taxon:10116"					
/sex="male"					
/tissue_type="liver"					
/dev_stage="adult"					
ORIGIN					
Query Match	50.2% Score 212.4; DB 10; Length 1181;				
Query Local Similarity	99.5%; Pred. No. 1.2e-48; Indels 0; Gaps 0;				
Matches	213; Conservative 0; Mismatches 1;				
QY	205 TCACAGCAAAACAACTTATTGACACGCGGATCTTACACGCTCCCTGCACAAATCA	264			
DB	823 TCACAGCAAAACAACTTATTGACACGCGGATCTTACACGCTCCCTGCACAAATCA	882			
QY	265 TTAACCCGTGCTCCGACGCGCCCTTATTAAGGCGGTATGCGCCAGCCAGCATGCT	324			
DB	883 TTAACCCGTGCTCCGACGCGCCCTTATTAAGGCGGTATGCGCCAGCCAGCATGCT	942			
QY	325 CCACGCGCGCGGACGACCAACCGAGCATTTGAACATGACACGCGCATGTGCC	384			

Db 943 CCACGCCCCGAGAGACACAAACCCAGCAGATTGACACACTGCAACGCGCATCTGCCC 1002
 Oy 385 AGAGAGCTTGAGCACCACTTCGGCTACTAGCTA 418
 Db 1003 AAGAGACTGTGACACCACTTCGGCTACTATCTA 1036

RESULT 2
 AC136382
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-97018, *** SEQUENCING IN PROGRESS
 *** 63 unordered pieces.
 AC136382
 AC136382.1 GI:24462257
 HTG: HTGS PHASE1.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS Muzny, D. Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levon, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshaw, L., Louieged, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M., McNeill, T., Meenan, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G., Olariunsegun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindere, A., Popovic, D., Primus, E., Pu, L., L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smales, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Wooley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weis, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G. and Gibbs, R. A.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Rat Genome Sequencing Consortium.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (01-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

 Center: Baylor College of Medicine
 Genome Center
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: KOGS
 Center clone name: CH230-97018

 Summary Statistics
 Sequencing vector: plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 135613 bases at least Q40
 Consensus quality: 140849 bases at least Q30
 Consensus quality: 145680 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a working draft sequence. It currently
 * consists of 63 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1010: contig of 1010 bp in length
 1011 1100: contig of 1498 bp in length
 2609 2708: gap of unknown length
 2709 3915: contig of 1207 bp in length
 3916 4015: gap of unknown length
 4016 5547: contig of 1532 bp in length
 5548 5647: gap of unknown length
 5648 7119: contig of 1472 bp in length
 7120 7219: gap of unknown length
 7220 8885: contig of 1666 bp in length
 8886 8985: gap of unknown length
 8986 10258: contig of 1273 bp in length
 10259 10358: gap of unknown length
 10359 11626: contig of 1268 bp in length
 11627 11726: gap of unknown length
 11727 13488: contig of 1762 bp in length
 13489 13588: gap of unknown length
 13589 15122: contig of 1534 bp in length
 15123 15222: gap of unknown length
 15223 16946: contig of 1724 bp in length
 16947 17047: gap of unknown length
 17048 18164: contig of 1118 bp in length
 18165 18264: gap of unknown length
 18265 19678: contig of 1414 bp in length
 19679 21687: gap of unknown length
 21688 21787: contig of 1909 bp in length
 21788 22892: gap of unknown length
 22893 22992: contig of 1105 bp in length
 22993 24336: contig of 1344 bp in length
 24337 24436: gap of unknown length
 24437 26517: contig of 2181 bp in length
 26518 26577: gap of unknown length
 26578 28577: contig of 1860 bp in length
 28578 30823: gap of unknown length
 30824 30923: gap of unknown length
 30924 32266: contig of 1343 bp in length
 32267 32366: gap of unknown length
 32367 33894: contig of 1528 bp in length
 33895 35373: gap of unknown length
 35374: contig of 1379 bp in length


```

* 35374 35473: gap of unknown length
* 35474 37295: contig of 1822 bp in length
* 37296 37395: gap of unknown length
* 37396 38420: contig of 1025 bp in length
* 38421 38520: gap of unknown length
* 38521 39974: contig of 1454 bp in length
* 39975 40074: gap of unknown length
* 40075 42283: contig of 2209 bp in length
* 42284 42383: gap of unknown length
* 42384 43721: contig of 1338 bp in length
* 43722 43821: gap of unknown length
* 43822 46538: contig of 2717 bp in length
* 46539 46638: gap of unknown length
* 46639 48621: contig of 1983 bp in length
* 48622 48721: gap of unknown length
* 48722 50602: contig of 1881 bp in length
* 50603 50702: gap of unknown length
* 50703 53841: contig of 3139 bp in length
* 53842 53942: gap of unknown length
* 53942 56195: contig of 2254 bp in length
* 56196 56295: gap of unknown length
* 56296 58949: contig of 2654 bp in length
* 58950 59049: gap of unknown length
* 59050 61032: contig of 1983 bp in length
* 61033 61132: gap of unknown length
* 61133 62942: contig of 1810 bp in length
* 62943 63042: gap of unknown length
* 63043 66045: contig of 3003 bp in length
* 66046 66145: gap of unknown length
* 66146 68197: contig of 2052 bp in length
* 68198 68297: gap of unknown length
* 68298 72235: contig of 3938 bp in length
* 72236 72335: gap of unknown length
* 72336 75803: contig of 3468 bp in length
* 75804 75903: gap of unknown length
* 75904 78355: contig of 2452 bp in length
* 78356 78455: gap of unknown length
* 78456 81458: contig of 3003 bp in length
* 81459 81558: gap of unknown length
* 81559 84673: contig of 3115 bp in length
* 84674 84773: gap of unknown length
* 84774 87739: contig of 2966 bp in length
* 87740 87839: gap of unknown length
* 87840 91535: contig of 3696 bp in length
* 91536 91635: gap of unknown length
* 91636 95522: contig of 3887 bp in length
* 95523 95622: gap of unknown length
* 95623 98730: contig of 3108 bp in length
* 98731 98830: gap of unknown length
* 98831 101618: contig of 2788 bp in length
* 101619 101718: gap of unknown length
* 101719 105535: contig of 3817 bp in length
* 105536 105635: gap of unknown length
* 105636 109393: contig of 3758 bp in length
* 109394 109493: gap of unknown length
* 109494 113564: contig of 4071 bp in length
* 113565 113664: gap of unknown length
* 113665 117630: contig of 3966 bp in length
* 117631 117730: gap of unknown length
* 117731 122502: contig of 4772 bp in length
* 122503 122602: gap of unknown length
* 122603 126644: contig of 4042 bp in length
* 126645 126744: gap of unknown length

```

Query Match 50.2%; Score 212.4; DB 2; Length 185148;
 Best Local Similarity 99.5%; Pred. No. 1,le=48;
 Matches 213; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 205 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTGAGACGCTGCGCTGACAAATCA 264
 DB 99253 TCACAAAGCAAAACAACTTATTGTAACACGGGGATCTGAGACGCTGCGCTGACAAATCA 99312
 QY 265 TTAACCGGTGCTGCGAGCCAGCCCTTCAATTAAGCCCTGGGTATGCGCCAGCAAGATGCT 324

```

DB 99313 TTAACCGGTGCTGCGAGCCAGCCCTTCAATTAAGCCCTGGGTATGCGCCAGCAAGATGCT 99372
QY 325 CCATGCCCCCGGAGACAAACCCAGGAGACATTGAACATCTGCACACGCGCATTCGCC 384
DB 99373 CCATGCCCCCGGAGACAAACCCAGGAGACATTGAACATCTGCACACGCGCATTCGCC 99432
QY 385 AGAGAGCTGTGACCAACCACTTCGGTACTAGCTA 418
DB 99433 AGAGAGCTGTGACCAACCACTTCGGTACTAGCTA 99466

RESULT 3
MUTIGF
LOCUS 1363 bp DNA linear ROD 01-AUG-1996
DEFINITION M.musculus gene for insulin-like growth factor binding protein-1.
ACCESSION X67493.1
VERSION X67493.1 GI:52659
KEYWORDS insulin-like growth factor binding protein-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1363)
AUTHORS Lee,J., Greenbaum,L., Haber,B.A., Nagle,D., Lee,V., Miles,V.,
Mohn,K.L., Bucan,M. and Taub,R.
TITLE Structure and localization of the IGFBP-1 gene and its expression
during liver regeneration
Hepatology 19 (3), 656-665 (1994)
MEDLINE 9416468
PUBMED 7509771
REFERENCE 2 (bases 1 to 1363)
AUTHORS Mohn,K.L., Waddie,J.R. and Taub,R.
TITLE Comparison of mouse and human IGFBP-1 genes reveals a potential
insulin-responsive sequence and conservation of all intron/exon
boundaries
Nucleic Acids Res.
REFERENCE 3 (bases 1 to 1363)
AUTHORS Taub,R.A.
TITLE Direct Submission
Submitted (23-JUL-1992) R.A. Taub, Univ. of Pennsylvania, Howard
Hughes Medical Institute, Clinical Research Bldg., Room 475, 422
Curie Boulevard, Philadelphia, PA 19104-6145, USA
FEATURES
source
1..1363
location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="UM2.1"
/cell_line="NIH 3T3"
/cell_type="fibroblast"
/clone_lib="genomic, lambda FIX II"
/dev_stage="embryo"
500..508
note="AP-2 consensus site"
692..706
note="insulin-responsive element"
720..741
note="Caat box with APF, HNF and NF-E1 consensus
sequences"
764..768
792..1336
number=1
792
note="mRNA cap site"
964..1336
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="CAA47832.1"
/db_xref="GI:52700"
/db_xref="GOA:P47876"
/db_xref="UniProt/Swiss-Prot:P47876"
misc_feature
misc_feature
misc_feature
CAAT_signal
TATA_signal
exon
misc_feature
CDS

```

ORIGIN

Query Match	44.5%	Score 188.4;	DB 10;	Length 1363;
Best Local Similarity	92.5%;	Pred. NO. 6.1e-42;		
Matches 198; Conservative	0;	Mismatches 16;	Indels 0;	Gaps 0;

QY	205	TCA	CAG	CAAA	CAAA	CTT	ATT	TGA	CA	CGG	GGT	CTT	CG	CA	CGC	CTG	CA	CA	TCA	264
Db	677	TCA	CAG	CAAA	CAAA	CTT <td>ATT <td>TGA</td> <td>CA</td> <td>CGG</td> <td>GGT</td> <td>CTT <td>CG</td> <td>CA</td> <td>CGC</td> <td>CTG <td>CA</td> <td>CA</td> <td>TCA <td>736</td> </td></td></td></td>	ATT <td>TGA</td> <td>CA</td> <td>CGG</td> <td>GGT</td> <td>CTT <td>CG</td> <td>CA</td> <td>CGC</td> <td>CTG <td>CA</td> <td>CA</td> <td>TCA <td>736</td> </td></td></td>	TGA	CA	CGG	GGT	CTT <td>CG</td> <td>CA</td> <td>CGC</td> <td>CTG <td>CA</td> <td>CA</td> <td>TCA <td>736</td> </td></td>	CG	CA	CGC	CTG <td>CA</td> <td>CA</td> <td>TCA <td>736</td> </td>	CA	CA	TCA <td>736</td>	736

QY 265 TTAAACCCGCTGCAGAGCAGCCCTTCAAAAGCCCTGGGTATGCGCAGCAGCATGGT 324

Db 737 TTAACTGTGCCGACAGCCAGCCCTTCAATAGGCTCTGGGTACGACCAAGCAGCATGGT 796

QY 325 CCACAGCCCCCGAGACACAAACCCAGCGGACATTGAACACTGCACACGGGCATCTGCC 384

Db 797 CCACTGCCCGCGGAGACACACACCCAGCGGACATTGAACACTGCACACGGGCCTGTGCC 856

QY 385 AGAGAGCTGTGATCCACCACCTTCGCTACTAGCTA 418
 DB 857 AGAGAGCTGTGATCCACCACCTTCGCTACTACTCTA 890

RESULT 4	
AL607124	
AL607124	199943 bp
	DNM
	140000
	PCD 11 100-2000

DEFINITION	mouse DNA sequence	LOC101928110	on chromosome 11, complete
ACCESSION	AL607124	GI:20145926	
VERSION	AL607124.15		

SOURCE ORGANISM	tax.	musculus (house mouse)	musculus	metazoa: chordata: craniata: euteleostomi: eukaryota: metazoa: chordata: vertebrata: euteleostomi: musculus
musculus (house mouse)	musculus	musculus	musculus	musculus

[illegible]

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On Apr 12, 2002 this sequence version replaced gi:19847866.

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission

only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMPEP; information on the WORMPEP

RP23-20C9 is from the RPC1-23 Mouse PAC library constructed by the group of Pieter de Jong.

```

VECTOR: PBACE3.6.
FEATURES
source      location/Qualifiers
1..192843

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"

```

```

/clone="RP23-20C9"
/clone_id="RPCL-23"

```

Query Match	44.5%	Score 188.4	DB 10	Length 192843
Best Local Similarity	92.5%	Pred. No. 5.7e-42		
Matches 198; Conservative	0	Mismatches 16	Indels 0	Gaps 0

Db 2429 TCACAGCAAAACAACCTATTTTTGAACACTGGGGTCTTAGACGCGTCCCTGACATCA 2488

Db 2489 TTACCTGTGCGCACAGCCAGCCCTTCATAAGGCTCTGGGTACGACCCAGCCAGCATGGT 2548

Db 2549 CCACTGCCCGCGAGACACACACCCAGCAGCATTTGAACATCGCACACGGCGCTGCCC 2608

Db 2609 AGAGAGCTGTGACCAACCATTTGCCACTATATCTA 2642

LOCUS	LENGTH	FEATURES	DATE
BC013345	1569 bp	mRNA	linear
BC013345			ROD 30-JUN-2004

ACCESSION (CUNA CLONE MGC:14075 IMAGE:416189), complete cas.
BC013345
VERSION BC013345.1 GI:15426482
REVIEWS
MGC

ORGANISM
Mus musculus (mouse)
Mus musculus (mouse)
Eukaryota; Metazoa: Chordata: Vertebrata; Euteleostomi; Mammalia: Eutheria: Rodentia: Scuriromorpha: Muridae: Murinae: Mus

114685 L.C. 1587
 Authors
 Struhsberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shemen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.

Diachenko, L., Matrusina, K., Farmer, A. A., Rudin, G. M., Hong, J., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavart, T. L., Scheetz, T. E., Brownstein, M. J., Uedín, T. B., Toshiyuki, S.

Abdramson, R.D., Mulahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.C., Malek, J.A., Gunnarone, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultyk, S.W., Hale, S.

Fahy, J., Helton, E., Kerteman, M., Maden, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D.

TITLE Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schmerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED
12477932
2 (bases 1 to 1569)
REFERENCE

TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK	COMMENT
USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk.	

Tissue Procurement: Jeffrey B. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org

contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Keltman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN.ac: http://image.llnl.gov
Series: IRK Plate: 18 Row: 1 Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

Location/Qualifiers

1..1569
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:14075 IMAGE:4161889"
/feature_type="liver, normal. 5 month old male mouse."
/clone_id="NCI CGAP_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1569
/gene="Igfbp1"
/note="synonym: IGFBP-1"
/db_xref="LocustID:16006"
/db_xref="MGI:96436"
204..1022
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH13945.1"
/db_xref="GI:15426483"
/db_xref="MGI:96436"
/db_xref="LocustID:16006"
/translation="MPEPLTVSWPLILLSPQIVAAAGAPQPHCAPCTARLGLCP
PVPASCPETSRPAGCCCTCALPMAAGVATATACAGLSCRALPGBRPALATRG
OGACVPEPAPATSTLSSQHEARAAVSADELSPEMTESQLDSHLMASRED
OPIIMNAISTVSMRAREMADIKKWKQPCRELKYLRLAAQKAGBEIYKFLPN
CNKGFYHSKQETSIDSEALCMCVPMGSRIPGSLERDPRCHQFVNVQN"

CDS

23.2%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 1.2e-16;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

ORIGIN

Query Match 23.2%; Score 98.2; DB 10; Length 1569;
Best Local Similarity 92.8%; Pred. No. 1.2e-16;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

RESULT 6
AY560836 1507 bp mRNA linear ROD 22-MAR-2004
LOCUS Spemophilus tridecemlineatus insulin-like growth factor binding
DEFINITION protein 1 (IGFBP1) mRNA, complete cds.
ACCESSION AY560836
VERSION AY560836.1 GI:45505308
KEYWORDS
SOURCE .
ORGANISM Spemophilus tridecemlineatus (thirteen-lined ground squirrel)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
Spemophilus.

REFERENCE 1 (bases 1 to 1507)
AUTHORS Li Y., Klimans D. and Hallenbeck J.M.
TITLE Cloning and characterization of insulin-like growth factor binding protein 1 (IGFBP-1) from thirteen-lined ground squirrel

JOURNAL

Unpublished
2 (bases 1 to 1507)

REFERENCE Li Y., Klimans D. and Hallenbeck J.M.
AUTHORS Direct Submission
TITLE Submitted (27-FEB-2004) StrokeBranch, NINDS/NIH, 36 Convent Dr.,
Bethesda, MD 20892, USA

FEATURES

Location/Qualifiers

1..1507
/organism="Spemophilus tridecemlineatus"
/mol_type="mRNA"
/db_xref="taxon:43179"
1..1507
/gene="IGFBP1"
689..1507
/gene="IGFBP1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAS67029.1"
/db_xref="GI:45505308"
/translation="MPEVPAPGLMPFLLLVAVYSTVASSTOPPHCAPCSAEKALCP
FVPSCEPLSRPAGCCCPWALPLGAAGVATATATAGLSCRALPGBRPALATRG
OGACVPEPAPATSTLSSQHEARAAVSADELSPEMTESQLDSHLMASSED
OPIIMNAISTVSMRAREMADIKKWKQPCRELKYLRLAAQKAGBEIYKFLPN
CNKGFYHSKQETSIDSEALCMCVPMGSRIPGSLERDPRCHQFVNVQN"

CDS

Query Match 18.1%; Score 76.6; DB 10; Length 1507;
Best Local Similarity 67.6%; Pred. No. 1.3e-10;
Matches 144; Conservative 0; Mismatches 54; Indels 15; Gaps 2;

ORIGIN

Query 204 ATCAACAAGCAAAACAACTTTATTTTGAACAGCGGATCTTAGACGCGCTGACCAATC 263
Db 429 AGCAGAGCAAAACAACTTTATTTTGAACAGCGGATCTTAGACGCGCTGACCAATC 488
Qy 264 ATTA-----CCGTCGTCCGAGCCAGCCCTCATAGGCGCTGAGTATGCCA 313
Db 489 ATTAATCTTCTAGTCAAGTGAAGTGAACCGCGCTTATTAAGCAAGGCTGACCCA 548
Qy 314 GCCAGCATGTCTCACTGCGCGCCGAGACAAACCCAGCAGCATTTGAACATGACACG 373
Db 549 GCAAGCATGTGCACTGCGCGCCGAGACAAACCCAGCAGCATTTGAACATTTGCACTTCA-----GCCACT 603
Qy 374 GCCATCTGCCAGAGAGCTGTGACCAACCACTTC 406
Db 604 GCCACTGCCAGAGAGCATTTGACCAACCACTTC 636

RESULT 7
AX401932 1500 bp DNA linear PAT 06-JUN-2002
LOCUS Sequence 1608 from Patent WO0210453.
DEFINITION AX401932
ACCESSION AX401932
VERSION AX401932.1 GI:21338112
KEYWORDS
SOURCE .
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1
AUTHORS Mendrick D., Porter M.W., Johnson K.R., Caetle A.L. and
Elashoff M.R.
TITLE Molecular toxicology modeling
JOURNAL Patent: WO 0210453-A 1608 07-FEB-2002;
Gene Logic, Inc. (US)

FEATURES

Location/Qualifiers

1..1500
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
/note="EMBL/GenBank Accession No. NM_013144"

ORIGIN

Query Match	18.1%;	Score 76.4;	DB 6;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 1.5e-10;		
Matches	77;	Conservative	0;	Mismatches 1;
Indels				
Gaps				
QY	341	CACAAACCAGGACGACATTGAAACATGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	400	
DB	8	CACAAACCAGGACGACATTGAAACATGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	67	
QY	401	CACCTTCGGCTACTACTA	418	
DB	68	CACCTTCGGCTACTACTA	85	
RESULT 8				
LOCUS	AX827271	1500 bp	DNA	linear
DEFINITION	Sequence 5 from Patent EP1344834.			
AX827271				
VERSION	AX827271.1	GI:39837360		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS	Boess,F., Suter-Dick,L. and Wolf,D.			
TITLE	Methods for the toxicity prediction of a compound			
JOURNAL	Patent: EP 1344834-A 5 17-SEP-2003;			
FEATURES	F. HOFMANN-LA ROCHE AG (CH)			
source	Location/Qualifiers			
	1..1500			
	/organism="Rattus norvegicus"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:10116"			
ORIGIN				
Query Match	18.1%;	Score 76.4;	DB 6;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 1.5e-10;		
Matches	77;	Conservative	0;	Mismatches 1;
Indels				
Gaps				
QY	341	CACAAACCAGGACGACATTGAAACATGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	400	
DB	8	CACAAACCAGGACGACATTGAAACATGTGCACACGCGCCATCTGCCAGAGAGCTGTGACAC	67	
QY	401	CACCTTCGGCTACTACTA	418	
DB	68	CACCTTCGGCTACTACTA	85	
RESULT 9				
LOCUS	RAT1GFB	1500 bp	mRNA	linear
DEFINITION	Rat 1GFB binding protein-1 (r1GFBP-1) mRNA, complete cds.			
ACCSSION	M56634			
VERSION	M56634.1	GI:204732		
KEYWORDS	IGF binding protein-1.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Rattus norvegicus			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
	1 (bases 1 to 1500)			
REFERENCE				
AUTHORS	Mohm,K.L., Melby,A.E., Tewari,D.S., Laz,T.M. and Taub,R.			
TITLE	The gene encoding rat insulinlike growth factor-binding protein 1			
JOURNAL	is rapidly and highly induced in regenerating liver			
MEDLINE	Mol. Cell. Biol. 11 (3), 1393-1401 (1991)			
COMMENT	Original			
FEATURES	source text: Rat, cDNA to mRNA.			
	Location/Qualifiers			
	1..1500			
	/organism="Rattus norvegicus"			

```

CDS
    /mol_type="mRNA"
    /strain="Fisher"
    /db_xref="taxon:10116"
    /tissue_type="regenerating liver"
    /dev_stage="adult"
    160..978
        /codon_start=1
        /product="IGF binding protein-1"
        /protein_id="AA41380.1"
        /db_xref="GI:204733"
        /translation="MPEELIVSWPFLILSFQVRVVGAPQPMHCACTAERLELCE
        FVPSPSEIRSPAGCCPTCALPIGAACGVATRCAGLSGRALPGPRPLHALTRQ
        GCAVLVBAAIPATSLSGSOHEEAKAVASDELAESEMEBOILHSFHIMAPSRREI
        OPTIMALISTSMRAREITDCLKKKEPCORELYKVLERLLAAOQKADELYKFTLP
        CNKKGFHSKOCETSLDGEGLCMCVIPWSGSKLPGSLERIGDPNCHYENVQN"

sig_peptide
mat_peptide
/product="IGF binding protein-1"

ORIGIN
Query Match      18.1%; Score 76.4; DB 10; Length 1500;
Best Local Similarity 98.7%; Pred. No. 1.5e-10;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCAGCGAGCATTTGAACACTGCACACGGCCATTGTGCCAGAGAGCTGTGACAC 400
|||||
DB 1 CACAAACCAGCGAGCATTTGAACACTGCACACGGCCATTGTGCCAGAGAGCTGTGACAC 60
|||||

QY 401 CACTTCGGCTACTAGCTA 418
|||||
DB 61 CACTTCGGCTACTATCTA 78
|||||

RESULT 11
RATIGFBA
```

LOCUS RATTIGFBA 5001 bp DNA linear ROD 30-NOV-1995
 DEFINITION Rat insulin-like growth factor binding protein-1 (IGFBP-1) gene,
 complete cds.
 ACCESSION L22979
 VERSION L22979.1 GI:1098472
 KEYWORDS insulin-like growth factor binding protein-1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 5001)
 AUTHORS Iacson, R., Oehler, D., Yang, B., Goewami, R. and Unterman, T.
 TITLE Bideoxy sequencing and structural analysis of the rat insulin-like
 growth factor binding protein-1 gene
 JOURNAL Biochim. Biophys. Acta 1218 (1), 95-98 (1994)
 MEDLINE 94250701
 PUBMED 7514892
 COMMENT On Nov 30, 1995 this sequence version replaced gi:185167.
 Original source text: Rattus norvegicus (strain Sprague-Dawley)
 DNA.

FEATURES
 source
 1. location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /sex="male"
 /tissue_type="liver"
 /dev_stage="adult"
 /tissue_lib="Clontech EMBL-3 Sp6/T7"
 60..525
 /number=1
 5'UTR 60..152
 gene 153..4221
 /gene="IGFBP-1"
 join(153..525,1850..2034,2874..3002,4090..4221)
 CDS /gene="IGFBP-1"
 /codon_start=1
 /product="insulin-like growth factor binding protein"
 /protein_id="AAA82581.1"
 /db_xref="GI:1098473"
 /translation="MPEPLTVSWPPLILISFQVAVVAGAPQPMHCAPCTAERLTC
 PVPASCPERISRPAGCGCCCTCALPGAAGCVATAACQSLSRALPGEPRPIHALTRG
 QACVLEPAAPATSSLSGSQHEAKAAVAASELVAEPEMTSEQLDSEPHLMAPERSD
 QPILMNAISTVSMRAREITDILKWKPEQRELYVLRLLAAQKADBEYKFLPN
 CNKGFYHSKOCETSLDGBAGLCWCYPMWGCKIDGSIETRGDPNCHQYFNVN"
 526..1849
 /gene="IGFBP-1"
 /number=1
 1850..2034
 /gene="IGFBP-1"
 /number=2
 2035..2873
 /gene="IGFBP-1"
 /number=2
 2874..3002
 /gene="IGFBP-1"
 /number=3
 3003..4089
 /gene="IGFBP-1"
 /number=3
 4090..4743
 /number=4
 4222..4743
 exon
 intron
 ORIGIN
 3'UTR
 Query Match 18.1%; Score 76.4; DB 10; Length 5001;
 Best Local Similarity 98.7%; Pred. No. 1.5e-10;
 Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

341 CACAAACCAGGAGCATTTGAACACTGCACGCGCATCTGCCAGAGAGCTGTGACCC 400
 |||||||

Db 1 CACAAACCAGGAGCATTTGAACACTGCACGCGCATCTGCCAGAGAGCTGTGACCC 60
 QY 401 CACTTCCGCTACTACTA 418
 Db 61 CACTTCCGCTACTACTA 78

RESULT 12
 BC078889 1510 bp mRNA linear ROD 03-AUG-2004
 LOCUS Rattus norvegicus insulin-like growth factor binding protein 1,
 mRNA (cDNA clone MGC:93595 IMAGE:7129185), complete cds.
 ACCESSION BC078889
 VERSION BC078889.1 GI:50927646
 KEYWORDS MGC.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1510)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshylyuk, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Hellon, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shchlenker, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smolins, D.E.,
 Schercher, A., Schein, J.E., Jones, S.J. and Merrit, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1510)
 DIRECTOR MGC Project.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 COMMENT
 REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing By: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil@stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 184 Row: e Column: 23
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6981079.
 Location/Qualifiers
 1..1510
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"

gene
/clone="MGC:93595 IMAGE:7129185"
/tissue_type="Kidney, rat (Brown Norway)"
/clone_id="NIN MGC_236"
/lab_host="DH10B"
/note="Vector: pExpress1"
1.1510
/gene="Igfbp1"
/note="synonyms: IGFBP, IGF-BP25, IBP1"
/db_xref="locusid:25685"
/db_xref="ratmap:44422"
/db_xref="RGD:2872"
142.960
/gene="Igfbp1"
/codon_start=1
/product="insulin-like growth factor binding protein 1"
/protein_id="AAH78889.1"
/db_xref="GI:50927647"
/db_xref="locusid:25685"
/db_xref="ratmap:44422"
/db_xref="RGD:2872"
/translation="MPEPLTVVSWPFLILSPQVRVVGAPQPMHCAPCTARLELCP
PYASCPHSRRAGCCCTCALPIGAAGVATACAGLSGRLPGRPRPLHATRG
QACVDEPAPATSSUSGQHEAKAAVSEDELAESEPMTEQLDSFILMAPERED
OPLMNAISTYSMPAREITDILKKMKEPCQRELYVLERLAAQKADDELYKFLPN
CKMGFYHSKQCEISLDGEAGLCWCVYPMWGKKIPGSIETREDPNCQHYFNVQN"

ORIGIN

Query Match 15.5%; Score 65.4; DB 10; Length 1510;
Best Local Similarity 98.5%; Pred. No. 1.8e-07;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 352 CGAGATTGACACGACGAGCCATCTGCCAGAGAGCTGTGACCACTTCGCGCTA 411
Db 1 CGAGATTGACACGACGAGCCATCTGCCAGAGAGCTGTGACCACTTCGCGCTA 60
412 CTAGCTA 418
61 CTATCTA 67

RESULT 13
AY095345 3886 bp DNA linear PRI 30-DEC-2002
LOCUS Papio anubis insulin-like growth factor binding protein-1 (IGFBP-1)
DEFINITION gene, partial cde.
ACCESSION AY095345
VERSION AY095345.1 GI:20853764
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
REFERENCE
AUTHORS Kim,J.-J., Taylor,H.S., Akbas,G.E., Foucher,I., Trembleau,A.,
Jaffe,R.C., Fazleabas,A.T. and Unterman,T.G.
TITLE Regulation of insulin-like growth factor binding protein-1 promoter
activity by FKHR and HOTA10 in primate endometrial cells
JOURNAL Biol. Reprod. 68 (1), 24-30 (2003)
PUBMED 12493691
REFERENCE
PUBMED 2 (bases 1 to 3886)
AUTHORS Kim,J.-J., Jaffe,R.C. and Fazleabas,A.T.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2002) Department of Obstetrics and Gynecology,
University of Illinois at Chicago, 820 S. Wood, Chicago, IL 60612,
USA

FEATURES
source
1.3886
location/Qualifiers
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
<3655. .>3886
/gene="IGFBP-1"

mRNA
<3655. .>3886
/gene="IGFBP-1"
/product="insulin-like growth factor binding protein-1"
3655. .>3886
/gene="IGFBP-1"
/codon_start=1
/product="insulin-like growth factor binding protein-1"
/protein_id="AAH73273.1"
/db_xref="GI:20853765"
/translation="MSEVPVAPRWLVLLITVQYGVATASAPWQCAPSAAETALCPV
PASCSEVTRSHAGCCCPMCAPIGAAGVATAR"

CDS

ORIGIN

Query Match 14.8%; Score 62.6; DB 9; Length 3886;
Best Local Similarity 64.0%; Pred. No. 1.1e-06;
Matches 135; Conservative 0; Mismatches 59; Indels 17; Gaps 2;

Qy 206 CACAAGCAAAACAACTTATTTTGAACGCGGATCTTAGACGCTGCTGACATCAT 265
Db 3366 CACTAGCAAAACAACTTATTTTGAACGCTGCTGACGCTGCGGCTGCCATCAT 3425
Qy 266 TAAACC-----GTGCTGCCGAGCAGCCCTTATAGGCGCTGGTATGCCAGC 315
Db 3426 TAACTCTCTGTGCAAGTGGCGCGCTGTGCTTATAGGCGCGCTGTGCCAGC 3485
Qy 316 CAGCATGTCCTCAGTCCCGCCGAGACACAAACCGAGAGATTGAACATGCACAGGC 375
Db 3486 AAGCATGCGCCACCGCCATCTTCATCCAGCAAGC-----ATCTGCCGCCGCGCGCCG 3538
Qy 376 CATCTGCCAGAGAGCTGTGACCACTTC 406
Db 3539 CACCTCCCAAGAGAGCACTGCCACCGCTCC 3569

RESULT 14
G67139 448 bp DNA linear STS 18-SEP-2000
LOCUS IGFBP1.X1.1 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G67139
ACCESSION G67139
VERSION G67139.1 GI:10186730
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Cox,D.G., Bolliot,C. and Canzian,F.
TITLE Genome Survey
JOURNAL Unpublished (2000)
COMMENT

Contact: Federico Canzian
Genome Analysis Group
International Agency for Research on Cancer
150, cours Albert-Thomas, F-69372 Lyon Cedex 08 France
Tel: +33-4-72738698
Fax: +33-4-72738388
Email: canzian@arc.fr
Primer A: TGCACTAGCAAAACAACCTT
Primer B: GGGTGACCTCGAGCAGC
STS size: 448
Protocol:
Template: 50 ng
Primer: 4 um each
dNTPs: 2 mM each
MgCl2: 1.5-2.5 mM
Tag: 0.05 units
Total Vol: 25 ul
Buffer:
MgCl2: 1.5-2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.4.

FEATURES
Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human"

STS
primer_bind
primer_bind
complement(431..448)
ORIGIN

Query Match 14.4%; Score 60.8; DB 11; Length 448;
Best Local Similarity 66.5%; Pred. No. 3.4e-06;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTTATTTTGAACAGGGGATCTTAGACAGCTGCCCTGACATCAT 265
DB 3 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGCGCGCTGCAATCAT 62
QY 266 TAACCC-----GTGCTGCCAGCGCCCTTATAGGCGCTGGGTATGGCCAGC 315
DB 63 TAACTCTCTGTCGCAAGTGCGCGCTGTGCTTATAGGTGCGCTGTGTCAGC 122
QY 316 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGAGCATTTGAACACTGC-ACACGG 374
DB 123 GAGCATCGGCAACCGCATTC-----CATCAGGAGCATCTGCCCGCGCGCGCG 174
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406
DB 175 CCACCTCCAGAGAGCACTGCGCACCGCTCC 206

RESULT 15

LOCUS AX409747 6128 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 2394 from Patent WO0229103.
ACCESSION AX409747
VERSION AX409747.1 GI:2142452
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Alvarez, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 2394 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
Location/Qualifiers
1..6128
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. M74587"

ORIGIN

Query Match 14.4%; Score 60.8; DB 6; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.4e-06;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTTATTTTGAACAGGGGATCTTAGACAGCTGCCCTGACATCAT 265
DB 473 CACTAGCAAAACAACTTATTTTGAACACTCAGCTCTAGCGCGCGCTGCAATCAT 532
QY 266 TAACCC-----GTGCTGCCAGCGCCCTTATAGGCGCTGGGTATGGCCAGC 315
DB 533 TAACTCTCTGTCGCAAGTGCGCGCTGTGCTTATAGGTGCGCTGTGTCAGC 592
QY 316 CAGCATGTCTCACTGCCCGCGAGACAAACCCAGAGCATTTGAACACTGC-ACACGG 374
DB 593 GAGCATCGGCAACCGCATTC-----CATCAGGAGCATCTGCCCGCGCGCGCG 644
QY 375 CCATCTGCCAGAGAGCTGTGACCACTTC 406

DB 645 CCACCTCCAGAGAGCACTGCGCACCGCTCC 676

Search completed: September 1, 2005, 03:27:00
Job time : 2637.79 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 20:18:35 ; Search time 3215.92 Seconds
(without alignments)
506.706 Million cell updates/sec

Title: US-09-972-916b-6

Perfect score: 423
Sequence: 1 catggcgagcgagcgagcgaccc.....ttccgctactagctagcgacg 423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gsa1.*
- 9: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.2	23.2	1013	2	BF236974
2	87.8	20.8	310	6	BY783538
3	87.8	20.8	310	6	BY783538
4	87.8	20.8	310	6	BY783538
5	86.8	20.5	401	1	AI196314
6	86.8	20.5	480	1	AI196314
7	86.8	20.5	706	1	AI196314
8	86.8	20.5	765	1	AI196314
9	86.8	20.5	785	1	AI196314
10	86.8	20.5	799	1	AI196314
11	86.8	20.5	811	6	CA478518
12	86.8	20.5	848	1	AI196314
13	86.8	20.5	852	1	AI196314
14	85.2	19.4	605	1	AI196314
15	82.2	19.4	360	1	AI196314
16	81.4	19.2	846	7	CO573026
17	80.4	19.0	692	7	CV127049
18	79.4	18.8	713	7	CO560662
19	75.4	17.8	748	7	CO575629
20	73.8	17.4	694	7	CV117001
21	73.2	17.3	332	7	W30013
22	72.4	17.1	615	2	BB660958
23	72.4	17.1	618	6	CD561711
24	72.4	17.1	632	1	AI892189

25	67.4	15.9	488	1	AA674302
26	65.4	15.5	801	7	CK472246
27	64.4	15.2	840	7	CK473709
28	62.8	14.8	759	1	AA105355
29	61.6	14.6	269	2	BB604790
30	61.6	14.6	587	2	AA916227
31	60.8	14.4	1589	3	CR621807
32	60.8	14.4	1601	3	CR595377
33	59	13.9	545	1	AA060360
34	55.8	13.2	597	8	AZ652514
35	55.2	13.0	1167	3	CR667848
36	53.6	12.7	1128	3	CR657703
37	53.6	12.7	1128	3	CR671810
38	53.6	12.7	1159	3	CR663312
39	50.4	11.9	289	2	BE066031
40	48.8	11.5	234	2	AA086850
41	48.8	11.5	325	2	AA085818
42	47.2	11.2	307	2	AA086846
43	47.2	11.2	659	8	AZ840793
44	46.2	10.9	212	2	BE149547
45	45.8	10.8	270	2	AA753361

ALIGNMENTS

RESULT 1
LOCUS BF236974 1013 bp mRNA linear EST 14-NOV-2000
DEFINITION 60202605F1.NCI.CGAP.L19 Mus musculus cDNA clone IMAGE:4161889 5',
RNA sequence.
ACCESSION BF236974.1 GI:11150891
VERSION BF236974.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS NIH-MGC
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contract: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM9443 row: j column: 02
High quality sequence stop: 581.
Location/Qualifiers
1..1013
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4161889"
/lab_host="DH10B (TI phase-resistant)"
/clone_lib="NCI CGAP L19"
/note="Organ: Liver; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 23.2%; Score 98.2; DB 2; Length 1013;
Best Local Similarity 92.8%; Pred. No. 3.7e-17;
Matches 103; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 308 TGGCAGCAGCATGTCTCCTCCCGCGAGACACAAACGAGCATTTGAACACTG 367

```

Db      9  TGGGCGAGGAGCATGTGTCACCTGCCCGGAGACACACACCGAGCATTTGAACACTG 68
QY      368  CACAGCGGCATTCGCCGAGAGCTGTGACCACTTCGCGACTAGCTA 418
        69  CACACGGCGCTGTGCCGAGAGAGCTGTGACCACTTCGCGACTAGCTA 119

RESULT 2
LOCUS   BY783538
DEFINITION  310 bp mRNA linear EST 23-MAR-2004
Mus musculus cDNA clone L930176D05 5', mRNA sequence.
ACCESSION  BY783538
VERSION    BY783538.1
KEYWORDS   GI:39710177
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M.,
AUTHORS    Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
            Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanuki, A.,
            Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
            Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
            Pavan, W., Aldinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T.,
            Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M.,
            Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
            Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
            Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
MEDLINE   22703353
PUBMED    12819125
COMMENT    Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES
source    Location/Qualifiers
            1..310
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="L930176D05"
            /tissue_type="whole body"
            /dev_stage="17.5 days embryo"
            /clone_idb="RIKEN full-length enriched, 17.5 days embryo
            whole body"

ORIGIN
Query Match      20.8%; Score 87.8; DB 6; Length 310;
Best Local Similarity 92.9%; Pred. No. 2.8e-14;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      320  ATGGTCCACTGCGCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 379
        2  ATGGTCCACTGCGCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 61
Db      2  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 418
QY      380  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 418

```

```

Db      62  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 100

RESULT 3
LOCUS   BY794229
DEFINITION  331 bp mRNA linear EST 23-MAR-2004
Mus musculus cDNA clone L930292H24 5', mRNA sequence.
ACCESSION  BY794229
VERSION    BY794229.1
KEYWORDS   GI:39720868
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  Carninci, P., Waki, K., Shiraki, T., Kono, H., Shibata, K., Itoh, M.,
AUTHORS    Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
            Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watanuki, A.,
            Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
            Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
            Pavan, W., Aldinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T.,
            Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagiolini, M.,
            Henrich, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
            Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
            Targeting a complex transcriptome: the construction of the mouse
            full-length cDNA encyclopedia
JOURNAL   Genome Res. 13 (6B), 1273-1289 (2003)
MEDLINE   22703353
PUBMED    12819125
COMMENT    Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center (GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.jp/) for
            further details.
FEATURES
source    Location/Qualifiers
            1..331
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6J"
            /db_xref="taxon:10090"
            /clone="L930292H24"
            /tissue_type="whole body"
            /dev_stage="17.5 days embryo"
            /clone_idb="RIKEN full-length enriched, 17.5 days embryo
            whole body"

ORIGIN
Query Match      20.8%; Score 87.8; DB 6; Length 331;
Best Local Similarity 92.9%; Pred. No. 2.8e-14;
Matches 92; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      320  ATGGTCCACTGCGCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 379
        2  ATGGTCCACTGCGCGCGAGACAAACCCAGGAGCATTTGAACACTGACACGGGCATC 61
Db      62  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 100
QY      380  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 418
        62  TGGCCAGAGAGCTGTGACCACTTCGCGTACTAGCTA 100

```

RESULT 4
BY71317

COMMENT

Contact: Yoshihide Hayaishi-zaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-resgsc.riken.jp URL: <http://genome.gsc.riken.jp/>
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.jp/>) for
further details.

FEATURES		Location/Qualifiers
source		1..354
		/organism="Mus musculus"
		/mol_type="mRNA"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/clone="L930052G15"
		/tissue_type="whole body"
		/dev_stage="17.5 days embryo"
		/clone_idb="RIKEN full-length enriched, 17.5 days embryo whole body"
ORIGIN		
Query Match	20.8%;	Score 87.8; DB 6; Length 354;
Best Local Similarity	92.9%;	Pred. No. 2.8e-14;
Matches	92; Conservative	0; Mismatches 7; Indels 0; Gaps 0
QY	320	ATGATCCATGTCCTCCGCGGAGACACAAACCCAGCAGCATTTGAACTGCACACGGCCATC 379
Db	2	ATGATCCATGTCCTCCGCGGAGACACAAACCCAGCAGCATTTGAACTGCACACGGCCATC 61
QY	380	TGCCCCAGAGAGCTGTGACCAACCACTTCGGCTACTGACTA 418
Db	62	TGCCCAAGAGAGCTGTGACCAACCACTTCGGCTACTGACTA 100

RESULT 5	LOCUS	DEFINITION	ACCESSION
A1785818	A1785818	401 bp. mRNA linear EST 02-JUL-1999 u178h05.y1 Sugano mouse liver m1s musculus cDNA clone IMAGE1288569 5' similar to gp:x81575 M.musculus mRNA for insulin-like growth factor binding (MOUSE) ;, mRNA sequence.	A1785818

VERSION	TITLE
KEYWORDS	JOURNAL
SOURCE	COMMENT
ORGANISM	
REFERENCE	
AUTHORS	
AT795818.1 GI:5333534 EST. <i>Mus musculus</i> (house mouse) <i>Mus musculus</i> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus 1 (bases 1 to 401)	Unpublished (1999) Other ESTs: u178n05.x1 Contract: Marra M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 850L, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watsn.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGS Consortium (info@image.llnl.gov) for further information.
GI:972893	
Seq primer: custom primer used High quality sequence stop: 126.	
Location/Qualifiers 1..401	

```

FEATURES
SOURCE
Location/Qualifiers
1. 401
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888569"
/sex="female"
/dev_stage="adult"
/lab_host="DRI08"
/clone_1b="Sugano mouse liver m1a"
/note="Organ: liver; Vector: pME185-FL3; Site 1: DraIII
(CACGTGTG) ; Site 2: DraIII (CACCATGTG) ; 1st strand cDNA
was primed with an oligo (dT) primer
[ATGTGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCCCTTACTG], digested
and cloned into distinct DraIII sites of the pME185-FL3
vector (5' site CACGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCCTGCTTAAAGTCTCG and 3' end
primer CGACCTGCAGCTCGACACAC."

```

ORIGIN	
Query Match	20.5%; Score 86.8; DB 1; Length 401;
Best Local Similarity	92.9%; Pred. No. 5,7e-14;
Matches	91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	321 TGGTCCACTGCGCGCGAGACACAAACCCAGCGAGCATTTGAACACTGTGCACACGGGCATCT 380
Db	1 TGGTCCACTGCGCGCGAGACACACACCCAGCGAGCATTTGAACACTGTGCACACGGCGCTCT 60
QY	381 GCCCAGAGAGCTGTGACCAACCATCTTCCGCTAACTAGCTA 418
Db	61 GCCCAGAGAGCTGTGACCAACCATTTGCCACTAATCTA 98

RESULT 6	
A1196314	
LOCUS	
DEFINITION	
A1196314 480 bp mRNA linear EST 14-OCT-1998	
u712407.y1 Sugano mouse liver m1.a Mus musculus cDNA clone	
IMAGE:1887828.5' similar to gb:A59316.tna1 INSULIN-LIKE GROWTH	
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN) : gb:x61579.m.musculus	
mRNA for insulin-like growth factor binding (MOUSE); mRNA	
sequence.	

ACCESSION A1196314 GI:3748920
 VERSION A1196314.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 480)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Schellberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Watson, R., and Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGS Consortium (info@image.llnl.gov) for further information.
 MGI:972152
 Seg primer: custom primer used
 High quality sequence stop: 375.
 Location/Qualifiers
 1..480
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1887828"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCGCTCGACACA."
 ORIGIN
 Query Match 20.5%; Score 86.8; DB 1; Length 480;
 Best Local Similarity 92.9%; Pred. No. 5,9e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 321 TGTGTCACCTGCCCCGAGACACAAACCAGAGAGATGAACACTGCACAGGCGACT 380
 Db 1 TGTGTCACCTGCCCCGAGACACACACCCAGAGAGATTGAACACTGCACAGGCGCT 60
 Oy 381 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 418
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 7
 A1530146 706 bp mRNA linear EST 18-MAR-1999
 LOCUS A1530146
 DEFINITION ui89f09.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1889609 5' similar to gb:M5316.mai INSULIN-LIKE GROWTH
 FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579.M.musculus
 mRNA for insulin-like growth factor binding (MOUSE);, mRNA

sequence.
 ACCESSION A1530146 GI:4444281
 VERSION A1530146.1
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 706)
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
 TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGS Consortium (info@image.llnl.gov) for further information.
 MGI:973933
 Seg primer: custom primer used
 High quality sequence stop: 479.
 Location/Qualifiers
 1..706
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1889609"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mla"
 /note="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII (CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTTACTG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end primer CGACTGCGCTCGACACA."
 ORIGIN
 Query Match 20.5%; Score 86.8; DB 1; Length 706;
 Best Local Similarity 92.9%; Pred. No. 6.4e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Oy 321 TGTGTCACCTGCCCCGAGACACAAACCAGAGAGATGAACACTGCACAGGCGACT 380
 Db 1 TGTGTCACCTGCCCCGAGACACACACCCAGAGAGATTGAACACTGCACAGGCGCT 60
 Oy 381 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 418
 Db 61 GCCCAGAGAGCTGTGACCAACCACTTCGGCTACTAGCTA 98
 RESULT 8
 A1098594 765 bp mRNA linear EST 20-AUG-1998
 LOCUS A1098594
 DEFINITION ue31e07.y1 Sugano mouse liver mla Mus musculus cDNA clone
 IMAGE:1481988 5' similar to gb:X81579.M.musculus mRNA for
 insulin-like growth factor binding (MOUSE);, mRNA sequence.
 ACCESSION A1098594

VERSION A1098594.1 GI:3448119
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:930344
Seq primer: custom primer used
High quality sequence stop: 395.
Location/Qualifiers
1..765
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1481988"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMB18-F13; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG]; digested and cloned into distinct DraIII sites of the pMB18-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCTGTCTTAAAGCTGCG and 3' end primer CGACTGTGACCTCGAGCA."

ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 765;
Best Local Similarity 92.9%; Pred. No. 6.5e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTTCACCTGCCCGCGAGACACAAACCCAGGACATTGAACACTGCACAGGCGCATCT 380
DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGGACATTGAACACTGCACAGGCGCATCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 98

RESULT 9
A1530313
LOCUS A1530313 785 bp mRNA linear EST 18-MAR-1999
DEFINITION u187c09.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1889785 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1530313
VERSION A1530313.1 GI:44444448
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 765)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974109
Seq primer: custom primer used
High quality sequence stop: 459.
Location/Qualifiers
1..785
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1889785"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pMB18-F13; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCCTACTG]; digested and cloned into distinct DraIII sites of the pMB18-F13 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTCTGTCTTAAAGCTGCG and 3' end primer CGACTGTGACCTCGAGCA."

ORIGIN
Query Match 20.5%; Score 86.8; DB 1; Length 785;
Best Local Similarity 92.9%; Pred. No. 6.6e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTTCACCTGCCCGCGAGACACAAACCCAGGACATTGAACACTGCACAGGCGCATCT 380
DB 1 TGTTCACCTGCCCGCGAGACACACACCCAGGACATTGAACACTGCACAGGCGCATCT 60
QY 381 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 418
DB 61 GCCCAGAGAGCTGTGACCAACCACTTCCGCTACTACTA 98

RESULT 10
A1529939
LOCUS A1529939 799 bp mRNA linear EST 18-MAR-1999
DEFINITION u187c09.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1889392 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M. musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION AI529939
 VERSION AI529939.1 GI:4444074
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 799)
 REFERENCE
 AUTHORS Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Maria M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:973716
 Seq primer: custom primer used
 High quality sequence stop: 506.
 FEATURES
 source
 Location/Qualifiers
 1..799
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1899392"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Sugano mouse liver mlia"
 /note="Organ: liver; Vector: pME188-FU3; Site 1: DraIII (CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer
 [ATGTGCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGCTGTGTG], digested and cloned into distinct DraIII sites of the pME188-FU3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTTGTCTCTTAAGCTGCG and 3' end primer CGACTCGACTCGAGCACACA."

ORIGIN

Query Match 20.5%; Score 86.8; DB 1; Length 799;
 Best Local Similarity 92.9%; Pred. No. 6.6e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTGTCACCTGCCCCGGAGACACAACCAGGACATTGAACACTGTGACACGGCCATCT 380
 DB 1 TGTGTCACCTGCCCCGGAGACACAACCAGGACATTGAACACTGTGACACGGCCGCT 60
 QY 381 GCCCAGAGAGCTGTGACCACTTCCTGCTACTACTA 418
 DB 61 GCCCAGAGAGCTGTGACCACTTCCTGCTACTACTA 98

RESULT 11
 LOCUS CA478518 811 bp mRNA linear EST 09-MAR-2004
 DEFINITION AGENCOURT 10789306 NIH MGC 152 Mus musculus cDNA clone
 IMAGE:6766752 5', mRNA sequence.
 ACCESSION CA478518
 VERSION CA478518.1 GI:24936188
 KEYWORDS EST.

SOURCE
 ORGANISM Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 811)
 REFERENCE
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittinger
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM00138 row: m column: 23
 High quality sequence stop: 536.
 FEATURES
 source
 Location/Qualifiers
 1..811
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6766752"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 152"
 /note="Organ: Liver; Vector: pDONR201; Site 1: attP2; Site 2: attP1; cDNA made by oligo-dT with attB2 site and directionally cloned. Priming sequence: 5'-TTTCGTGACGGCCGGACCACTTGTGAAGAAGCTGGGTTTTTTTTTTT TTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen, library amplification 16 cycles, library constructed by Mark Bittinger in the Bradfield Laboratory (McArdle Laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH-MGC Library."

ORIGIN

Query Match 20.5%; Score 86.8; DB 6; Length 811;
 Best Local Similarity 92.9%; Pred. No. 6.6e-14;
 Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGTGTCACCTGCCCCGGAGACACAACCAGGACATTGAACACTGTGACACGGCCATCT 380
 DB 39 TGTGTCACCTGCCCCGGAGACACAACCAGGACATTGAACACTGTGACACGGCCGCT 98
 QY 381 GCCCAGAGAGCTGTGACCACTTCCTGCTACTACTA 418
 DB 99 GCCCAGAGAGCTGTGACCACTTCCTGCTACTACTA 136

RESULT 12
 LOCUS A1790802 848 bp mRNA linear EST 02-JUL-1999
 DEFINITION uk28b10.y1 Sugano mouse kidney mlia Mus musculus cDNA clone
 IMAGE:1970299 5' similar to gb:X81579 M.musculus mRNA for insulin-like growth factor binding (MOUSE);, mRNA sequence.

ACCESSION A1790802
 VERSION A1790802.1 GI:5338518
 KEYWORDS EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 848)
 REFERENCE
 AUTHORS Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT

Other ESTs: uk28b10.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:987039

FEATURES

High quality sequence stop: 514.
Location/Qualifiers
1..848

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1970299"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse kidney mla"
/note="Organ: kidney; Vector: pME188-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME188-FL3
vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTAAAGCTGG and 3' end
primer CGACTGCAGCTCGAGCA."

ORIGIN

Query Match 20.5%; Score 86.8; DB 1; Length 848;
Best Local Similarity 92.9%; Pred. No. 6.7e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGGTCCACCTGCGCGGAGACACAAACCGAGAGCATTTGAACACTGCACAGCGGCATCT 380
|||||
Db 1 TGGTCCACCTGCGCGGAGACACACACCGAGAGCATTTGAACACTGCACAGCGGCCTCT 60
QY 381 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 418
|||||
Db 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 13
A1528304 852 bp mRNA linear EST 18-MAR-1999
LOCUS A1528304
DEFINITION u195g10.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1890210 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1528304 GI:4442439
VERSION A1528304
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 852)
REFERENCE Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Peyron,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schuck,R.,
Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

JOURNAL

COMMENT

Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:974534

FEATURES

High quality sequence stop: 478.
Location/Qualifiers
1..852

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1890210"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="Sugano mouse liver mla"
/note="Organ: liver; Vector: pME188-FL3; Site 1: DraIII
(CACTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTG], digested
and cloned into distinct DraIII sites of the pME188-FL3
vector (5' site CACTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCGTCTCTAAAGCTGG and 3' end
primer CGACTGCAGCTCGAGCA."

ORIGIN

Query Match 20.5%; Score 86.8; DB 1; Length 852;
Best Local Similarity 92.9%; Pred. No. 6.7e-14;
Matches 91; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 321 TGGTCCACCTGCGCGGAGACACAAACCGAGAGCATTTGAACACTGCACAGCGGCATCT 380
|||||
Db 1 TGGTCCACCTGCGCGGAGACACACACCGAGAGCATTTGAACACTGCACAGCGGCCTCT 60
QY 381 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 418
|||||
Db 61 GCCCAGAGAGCTGTGACCACTTCGCTACTAGCTA 98

RESULT 14
A1196154 605 bp mRNA linear EST 14-OCT-1998
LOCUS A1196154
DEFINITION u169d08.y1 Sugano mouse liver mla Mus musculus cDNA clone
IMAGE:1887663 5' similar to gb:M59316 rat INSULIN-LIKE GROWTH
FACTOR BINDING PROTEIN 1 PRECURSOR (HUMAN); gb:X81579 M.musculus
mRNA for insulin-like growth factor binding (MOUSE);, mRNA
sequence.

ACCESSION A1196154 GI:3748760
VERSION A1196154
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 605)
REFERENCE Marra,M., Hillier,L., Allen,M., Bowers,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMNI Mouse EST Project

JOURNAL
COMMENT

Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HM1 Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:971987

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 420.
Location/Qualifiers

```
1..605
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1887663"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mlia"
/notes="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCGCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
primer CGACTCTGCAGCTCGAGCACA."
```

ORIGIN

Query Match 20.1%; Score 85.2; DB 1; Length 605;
Best Local Similarity 91.8%; Pred. No. 1.8e-13;

Matches 90; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 321 TGGTCACTGCGCGCGAGACACAAACCGAGAGATTGAACACTGACACAGCGCATCT 380
|||||
1 TGGTCACTGCGCGCGAGACACACACCGAGAGATTGAACACTGACACAGCGCATCT 60

Oy 381 GCCCAGAGAGCTGTGACCACTTCGGCTACTACTCTA 418
|||||
61 GCCCAGAGAGCTGTGACCACTTCGGCTACTACTCTA 98

RESULT 15

A1785039

LOCUS 380 bp mRNA linear EST 02-JUL-1999
DEFINITION u173a06.y1 Sugano mouse liver mlia Mus musculus cDNA clone

IMAGE:1888018 5' similar to gp:X81579 M.musculus mRNA for
insulin-like growth factor binding (MOUSE); mRNA sequence.

ACCESSION

A1785039

VERSION

A1785039.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 380)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)

COMMENT

Other ESTs: u173a06.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watsn.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MG1:972342

FEATURES

source

Seq primer: custom primer used
High quality sequence stop: 353.
Location/Qualifiers

```
1..380
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1888018"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Sugano mouse liver mlia"
/notes="Organ: liver; Vector: pME18S-FL3; Site 1: DraIII
(CACTGTGTG); Site 2: DraIII (CACCATGTG); 1st strand cDNA
was primed with an oligo(dT) primer
[ATGTGCGCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCGCTACTG], digested
and cloned into distinct DraIII sites of the pME18S-FL3
vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
be used to isolate the cDNA insert. Size selection was
performed to exclude fragments <1.5kb. Library
constructed by Dr. Sumio Sugano (University of Tokyo
Institute of Medical Science). Custom primers for
sequencing: 5' end primer CTTCTGCTCTTAAAGCTGCG and 3' end
primer CGACTCTGCAGCTCGAGCACA."
```

ORIGIN

Query Match 19.4%; Score 82.2; DB 1; Length 380;
Best Local Similarity 91.6%; Pred. No. 1.2e-12;

Matches 87; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Oy 324 TCCACTGCGCGCGAGACACAAACCGAGAGATTGAACACTGACACAGCGCATCTGCC 383
|||||
1 TCCACTGCGCGCGAGACACACACCGAGAGATTGAACACTGACACAGCGCATCTGCC 60

Oy 384 CAGAGAGCTGTGACCACTTCGGCTACTACTCTA 418
|||||
61 CAGAGAGCTGTGACCACTTCGGCTACTACTCTA 95

Search completed: September 1, 2005, 06:56:53
Job time : 3217.09 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 03:27:12 ; Search time 1231.96 Seconds
 (without alignments)
 2248.600 Million cell updates/sec

Title: US-09-972-916B-6

Perfect score: 423
 Sequence: 1 catggggcagcaggggacac.....ttccgctactagctagcgcg 423

Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 327456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications NA: *
 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: *
 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: *
 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: *
 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: *
 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: *
 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq: *
 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq: *
 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq: *
 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq: *
 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq: *
 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq: *
 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq: *
 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq: *
 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	423	100.0	423	9	US-09-972-916A-6
2	270	63.8	270	9	US-09-972-916A-3
3	223.6	52.9	372	9	US-09-972-916A-5
4	219.4	51.9	321	9	US-09-972-916A-4
5	219	51.8	219	9	US-09-972-916A-2
6	154	36.4	372	9	US-09-972-916A-5
7	103	24.3	321	9	US-09-972-916A-4

8	76.4	18.1	1500	9	US-09-917-800A-1608	Sequence 1608, Ap
9	76.4	18.1	1500	17	US-10-388-934-5	Sequence 5, Appl
10	76.4	18.1	1500	17	US-10-191-803-73	Sequence 73, Appl
11	76.4	18.1	1500	18	US-10-152-319A-1613	Sequence 1613, Ap
12	60.8	14.4	6128	9	US-09-880-107-2393	Sequence 2393, Ap
13	60.8	14.4	6128	22	US-10-756-149-1484	Sequence 1484, Ap
14	60.8	14.4	9173	22	US-10-893-315-126	Sequence 126, App
15	60.8	14.4	9173	22	US-10-893-315-160	Sequence 160, App
16	55.2	13.0	13011	17	US-10-388-934-36	Sequence 36, Appl
17	55.2	13.0	13011	18	US-10-152-319A-2150	Sequence 2150, Ap
18	51	12.1	51	9	US-09-972-916A-1	Sequence 1, Appl1
19	48.4	11.4	3191	16	US-10-029-386-20306	Sequence 20306, A
20	46.2	10.9	12850	14	US-10-017-122-1	Sequence 1, Appl1
21	45.6	10.8	586	16	US-10-029-386-5186	Sequence 5186, Ap
22	43.6	10.3	90556	21	US-10-741-600-17972	Sequence 17972, A
23	42.8	10.1	423	9	US-09-972-916A-6	Sequence 6, Appl1
24	41	9.7	3037	20	US-10-363-345A-40659	Sequence 40659, A
25	41	9.7	3037	20	US-10-363-345A-40660	Sequence 40660, A
26	41	9.7	3037	21	US-10-363-483A-40659	Sequence 40659, A
27	41	9.7	3037	21	US-10-363-483A-40660	Sequence 40660, A
28	40.8	9.6	28000	16	US-10-091-625-11	Sequence 11, Appl
29	40.8	9.6	28000	17	US-10-096-399A-11	Sequence 11, Appl
30	40.8	9.6	28000	17	US-10-461-668-11	Sequence 11, Appl
31	40.8	9.6	28000	17	US-10-363-345A-40659	Sequence 381, App
32	40.8	9.6	148083	22	US-10-756-149-4571	Sequence 4571, Ap
33	40.2	9.5	31703	17	US-10-085-117-172	Sequence 17, App
34	39.4	9.3	201	21	US-10-741-600-69952	Sequence 69952, A
35	39	9.2	201	20	US-10-719-993-35505	Sequence 35505, A
36	38.6	9.1	34217	20	US-10-719-993-6865	Sequence 6865, Ap
37	38.4	9.1	11009	9	US-09-845-583-1	Sequence 1, Appl1
38	38.4	9.1	11009	14	US-10-037-182-3	Sequence 3, Appl1
39	38.4	9.1	11009	21	US-10-764-420-947	Sequence 947, App
40	37.8	8.9	201	20	US-10-719-993-35504	Sequence 35504, A
41	37.8	8.9	201	20	US-10-719-993-35509	Sequence 35509, A
42	37.8	8.9	905	20	US-10-363-345A-40663	Sequence 40663, A
43	37.8	8.9	905	20	US-10-363-345A-40664	Sequence 40664, A
44	37.8	8.9	905	21	US-10-363-483A-40664	Sequence 40664, A
45	37.8	8.9	905	21	US-10-363-483A-40664	Sequence 40664, A

ALIGNMENTS

RESULT 1
 US-09-972-916A-6
 ; Sequence 6, Application US/09972916A
 ; Patent No. US20020107198A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Thule, Peter M.
 ; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
 ; FILE REFERENCE: US 1292/01 (VA)
 ; CURRENT APPLICATION NUMBER: US/09/972,916A
 ; CURRENT FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 60/239,113
 ; PRIOR FILING DATE: 2000-10-11
 ; NUMBER OF SEQ ID NOS: 6
 ; SEQ ID NO 6
 ; LENGTH: 423
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthesized
 ; US-09-972-916A-6

Query Match 100.0%; Score 423; DB 9; Length 423;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-116;
 Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CATGGGCGCAGCGGGGCACTCCCGTTCCTTGACTTGGCCCCAGTGTACATGAGCGGC 60
 DB 1 CATGGGCGCAGCGGGGCACTCCCGTTCCTTGACTTGGCCCCAGTGTACATGAGCGGC 60
 QY 61 ACGGGGCACTCCCGTTCCTTGACTTGGCCCCAGTGTACATGAGCGGCAC 120

```

Db      61  AGGGGCACTCCCGTGGTTCCTGACTGGCCCCCAGGTACATGGGCGCAGGGGCGAC 120
Qy      121  TCCCGGTGTTCTGGACTCTGGCCCCCAGGTATCATGGGCGACGGGGGACTCCCGTGT 180
Db      121  TCCCGGTGTTCTGGACTCTGGCCCCCAGGTATCATGGGCGACGGGGGACTCCCGTGT 180
Qy      181  TCCCTGACTCTGGCCCCCAGGTATCATGAAGAAAACAATTAATTTTGAACAGGGGAT 240
Db      181  TCCCTGACTCTGGCCCCCAGGTATCATGAAGAAAACAATTAATTTTGAACAGGGGAT 240
Qy      241  CCTAGACGCTGCTCCTGACATCATTAACCCGTGCTGCCAGCCGACCCCTTATTAAGCC 300
Db      241  CCTAGACGCTGCTCCTGACATCATTAACCCGTGCTGCCAGCCGACCCCTTATTAAGCC 300
Qy      301  CTGGGTATGGCAGCAGCATGTGTCACCTGCCGCCGAGACACAAACCCAGGAGCATTTG 360
Db      301  CTGGGTATGGCAGCAGCATGTGTCACCTGCCGCCGAGACACAAACCCAGGAGCATTTG 360
Qy      361  AACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCACTTCCGCTACTAGTAGC 420
Db      361  AACACTGCACACGGCCATCTGCCAGAGAGCTGTGACCACTTCCGCTACTAGTAGC 420
Qy      421  CGC 423
Db      421  CGC 423

```

RESULT 2

```

US-09-972-916A-3
; Sequence 3, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 3
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-3

```

Query Match 63.8%; Score 270; DB 9; Length 270;

Best Local Similarity 100.0%; Pred. No. 2.7e-83;

Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      154  CATGGGCGCAGCGGGCACTCCCGTGGTTCCTGGAATCTGGCCCCCAGGTATCACAAGCA 213
Db      1  CATGGGCGCAGCGGGCACTCCCGTGGTTCCTGGAATCTGGCCCCCAGGTATCACAAGCA 60
Qy      214  AAACAACTTAATTTGAACAAGGGATCTGACAGCTGCCCTGACATCATTAACCGGT 273
Db      61  AAACAACTTAATTTGAACAAGGGATCTGACAGCTGCCCTGACATCATTAACCGGT 120
Qy      274  GCTGCGAGCAGCGCCCTTCAATAGGCTCTGGGTATGGCAGCAGCATGATGTCATGCGCC 333
Db      121  GCTGCGAGCAGCGCCCTTCAATAGGCTCTGGGTATGGCAGCAGCATGATGTCATGCGCC 180
Qy      334  GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 393
Db      181  GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 240
Qy      394  TGACCACCACTTCGGCTACTAGTAGCGGC 423
Db      241  TGACCACCACTTCGGCTACTAGTAGCGGC 270

```

RESULT 3

```

US-09-972-916A-5
; Sequence 5, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized
US-09-972-916A-5

```

Query Match 52.9%; Score 223.6; DB 9; Length 372;

Best Local Similarity 76.2%; Pred. No. 3.9e-67;

Matches 297; Conservative 0; Mismatches 74; Indels 19; Gaps 1;

```

Qy      34  ACTGCGCCCGCAGGTACATGGGCGCAGGGGCACTCCCGTGGTTCCTGACTGTGGCC 93
Db      2  AACTGGGGGCGCAGAGTCCAGAACACAGGGAATGCCCGCGCCCATGTACACTGGGG 61
Qy      94  CCCAGTGTACATGGGCGCAGCGGCACTCCCGTGGTTCCTGGAATCTGGCCCCCAGGTA 153
Db      62  GCCAGAGTCCAGGAACACAGGGAATGCCCGCGCCCATGTACACTGGGGGCGCAGAGTC 121
Qy      154  CATGGGCGCAGCGGCACTCCCGTGGTTCCTGGAATCTGGCCCCCAGGTATCACAAGCA 213
Db      122  CAGAAACCAAGGAGTGGCCCGTGGCCCATG-----TCAACAAGCA 162
Qy      214  AAACAACTTAATTTGAACAAGGGATCTGACAGCTGCCCTGACATCATTAACCGGT 273
Db      163  AAACAACTTAATTTGAACAAGGGATCTGACAGCTGCCCTGACATCATTAACCGGT 222
Qy      274  GCTGCGAGCAGCGCCCTTCAATAGGCTCTGGGTATGGCAGCAGCATGATGTCATGCGCC 333
Db      223  GCTGCGAGCAGCGCCCTTCAATAGGCTCTGGGTATGGCAGCAGCATGATGTCATGCGCC 282
Qy      334  GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 393
Db      283  GCCGAGACAAACCAGGAGCATTTGAACACTGCACACGGCCATCTGCCAGAGAGCTG 342
Qy      394  TGACCACCACTTCGGCTACTAGTAGCGGC 423
Db      343  TGACCACCACTTCGGCTACTAGTAGCGGC 372

```

RESULT 4

```

US-09-972-916A-4
; Sequence 4, Application US/09972916A
; Patent No. US20020107198A1
; GENERAL INFORMATION:
; APPLICANT: Thule, Peter M.
; TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION
; FILE REFERENCE: US 1292/01 (VA)
; CURRENT APPLICATION NUMBER: US/09/972,916A
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/239,113
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 4
; LENGTH: 321
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 51.9%; Score 219.4; DB 9; Length 321;
Best Local Similarity 99.5%; Pred. No. 1.1e-65;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 TATCAGACAAACAACTATTATTGAAACAGGGGATCTTAGACAGCTGCCCTGACAT 262
DB 101 TGTCAACAAACAAACAACTATTATTGAAACAGGGGATCTTAGACAGCTGCCCTGACAT 160
QY 263 CATTAACCCGTGCTGCCGAGCCGACGAGCCCTTATAGGCTGGGTATGAGCCAGCAGATG 322
DB 161 CATTAACCCGTGCTGCCGAGCCGACGAGCCCTTATAGGCTGGGTATGAGCCAGCAGATG 220
QY 323 GTTCACTGCCCGCCGAGACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGC 382
DB 221 GTTCACTGCCCGCCGAGACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGC 280
QY 383 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 423
DB 281 CAGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 321

RESULT 5

US-09-972-916A-2
Sequence 2, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 2

LENGTH: 219

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

US-09-972-916A-2

Query Match

Best Local Similarity 100.0%; Pred. No. 1.3e-65;

Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 TCACAGACAAACAACTATTATTGAAACAGGGGATCTTAGACAGCTGCCCTGACATCA 264
DB 1 TCACAGACAAACAACTATTATTGAAACAGGGGATCTTAGACAGCTGCCCTGACATCA 60
QY 265 TTAACCCGTGCTGCCGAGCCGACGAGCCCTTATAGGCTGGGTATGAGCCAGCAGATGT 324
DB 61 TTAACCCGTGCTGCCGAGCCGACGAGCCCTTATAGGCTGGGTATGAGCCAGCAGATGT 120
QY 325 CCATGCCCCCGGAGACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCC 384
DB 121 CCATGCCCCCGGAGACAAACCCAGCAGCATTTGAACACTGCACACGGCCATCTGCC 180
QY 385 AGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 423
DB 181 AGAGAGCTGTGACCAACCACTTCGGCTACTAGTACGCGC 219

RESULT 6

US-09-972-916A-5/c
Sequence 5, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/239,113
PRIOR FILING DATE: 2000-10-11
NUMBER OF SEQ ID NOS: 6

SEQ ID NO 5
LENGTH: 372
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Synthesized
US-09-972-916A-5

Query Match 36.4%; Score 154; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 5.9e-43;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ACATGGGCGCACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGCG 110
DB 154 ACATGGGCGCACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGCG 95
QY 111 CACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 170
DB 94 CACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 35
QY 171 CTCCTGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 204
DB 34 CTCCTGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 1

RESULT 7

US-09-972-916A-4/c
Sequence 4, Application US/09972916A
Patent No. US20020107198A1

GENERAL INFORMATION:

APPLICANT: Thule, Peter M.

TITLE OF INVENTION: GLUCOSE SENSITIVE REGULATOR OF INSULIN TRANSCRIPTION

FILE REFERENCE: US 1292/01 (VA)

CURRENT APPLICATION NUMBER: US/09/972,916A

PRIOR FILING DATE: 2001-10-10

PRIOR APPLICATION NUMBER: US 60/239,113

PRIOR FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 6

SEQ ID NO 4

LENGTH: 321

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthesized
US-09-972-916A-4

Query Match 24.3%; Score 103; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 2.9e-25;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ACATGGGCGCACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGCG 110
DB 103 ACATGGGCGCACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGCG 44
QY 111 CACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 153
DB 43 CACGGGGCACTCCCGTGTCTGACTCTGGCCCCCAGTGTACATGGGGCGGCA 1

RESULT 8

US-09-917-800A-1608
Sequence 1608, Application US/09917800A
Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castle, Arthur

Query Match	18.1%;	Score 76.4;	DB 17;	Length 1500;
Best Local Similarity	98.7%;	Pred. No. 7.9e-16;		

RESULT 1
US-10-152-319A-1613
Sequence 1613, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-0589-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335

;; PRIOR FILING DATE: 2001-05-22
;; PRIOR APPLICATION NUMBER: US 60/297,523
;; PRIOR FILING DATE: 2001-06-13
;; PRIOR APPLICATION NUMBER: US 60/298,925
;; PRIOR FILING DATE: 2001-06-19
;; PRIOR APPLICATION NUMBER: US 60/303,810
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/303,807
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/303,808
;; PRIOR FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: US 60/315,047
;; PRIOR FILING DATE: 2001-08-28
;; PRIOR APPLICATION NUMBER: US 60/324,928
;; PRIOR FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: US 60/330,867
;; PRIOR FILING DATE: 2001-11-01
;; PRIOR APPLICATION NUMBER: US 60/330,462
;; PRIOR FILING DATE: 2001-10-22
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 1613
;; LENGTH: 1500
;; TYPE: DNA
;; ORGANISM: Rattus norvegicus
;; FEATURE:
;; OTHER INFORMATION: Genbank Accession No. NM_013144
US-10-152-319A-1613

Query Match 18.1%; Score 76.4; DB 18; Length 1500;
Best Local Similarity 98.7%; Pred. No. 7.9e-16;
Matches 77; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 341 CACAAACCCAGGAGCATTTGAACACTGCAACGGCCATTGCCAGAGAGCTGTACAC 400
DB 8 CACAAACCCAGGAGCATTTGAACACTGCAACGGCCATTGCCAGAGAGCTGTACAC 67

QY 401 CACTCCGCTACTAGCTA 418
DB 68 CACTCCGCTACTAGCTA 85

RESULT 12
US-09-880-107-2393
;; Sequence 2393, Application US/09880107
;; Patent No. US20020142981A1
;; GENERAL INFORMATION:
;; APPLICANT: Horne, Darci T.
;; APPLICANT: Vockley, Joseph G.
;; APPLICANT: Scherf, Uwe
;; APPLICANT: Gene Logic, Inc.
;; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;; FILE REFERENCE: 44921-5028-WO
;; CURRENT APPLICATION NUMBER: US/09/880,107
;; PRIOR FILING DATE: 2001-06-14
;; PRIOR APPLICATION NUMBER: US 60/211,379
;; PRIOR FILING DATE: 2000-06-14
;; PRIOR APPLICATION NUMBER: US 60/237,054
;; PRIOR FILING DATE: 2000-10-02
;; NUMBER OF SEQ ID NOS: 3950
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 2393
;; LENGTH: 6128
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M74587
US-09-880-107-2393

Query Match 14.4%; Score 60.8; DB 9; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAAGCAAAACAATTATTTGAACGSGGATCTAGACGCTGCCCTGACAATCAT 265
DB 473 CACTGCAAAACAATTATTTGAACACTAGCTCTAGCGTGGCGCTGCCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCAGCCCTTCAATAGGCCCTGGATAGCCAGC 315
DB 533 TAACCTCGGTGCAAGTGGCGGCGCTGTGCTTTAATAGTGGCGCTGTGTCAGC 592
QY 316 CAGCATGTCCACTGCGCGCGGAGACAAACCCAGGAGATTGAACATGCG-ACACGG 374
DB 533 GAGCATCGGCCACCCCATTC-----CATCCAGGAGCATGTGCGCGCGCGCGCG 644
QY 375 CCATCGCCAGAGAGCTGTGACCACTTC 406
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 13
US-10-756-149-1484
;; Sequence 1484, Application US/10756149
;; Publication No. US20050181375A1
;; GENERAL INFORMATION:
;; APPLICANT: Aziz, Natsasha
;; APPLICANT: Zlotnick, Albert
;; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;; FILE REFERENCE: file
;; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;; CURRENT APPLICATION NUMBER: US/10/756,149
;; PRIOR FILING DATE: 2004-01-12
;; NUMBER OF SEQ ID NOS: 5818
;; SOFTWARE: Patentln version 3.2
;; SEQ ID NO 1484
;; LENGTH: 6128
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-756-149-1484

Query Match 14.4%; Score 60.8; DB 22; Length 6128;
Best Local Similarity 66.5%; Pred. No. 3.1e-10;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAAGCAAAACAATTATTTGAACGSGGATCTAGACGCTGCCCTGACAATCAT 265
DB 473 CACTGCAAAACAATTATTTGAACACTAGCTCTAGCGTGGCGCTGCCAATCAT 532
QY 266 TAACCC-----GTGCTGCCGAGCAGCCCTTCAATAGGCCCTGGATAGCCAGC 315
DB 533 TAACCTCGGTGCAAGTGGCGGCGCTGTGCTTTAATAGTGGCGCTGTGTCAGC 592
QY 316 CAGCATGTCCACTGCGCGCGGAGACAAACCCAGGAGATTGAACATGCG-ACACGG 374
DB 593 GAGCATCGGCCACCCCATTC-----CATCCAGGAGCATGTGCGCGCGCGCGCG 644
QY 375 CCATCGCCAGAGAGCTGTGACCACTTC 406
DB 645 CCACCTCCAGAGAGCACTGGCCACCGCTCC 676

RESULT 14
US-10-893-315-126
;; Sequence 126, Application US/10893315
;; Publication No. US20050147987A1
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH TYPE II DIABETES AND OBESITY, METHODS OF DETECTION AND
;; FILE REFERENCE: CLO00786
;; CURRENT APPLICATION NUMBER: US/10/893,315
;; PRIOR FILING DATE: 2004-07-19
;; PRIOR APPLICATION NUMBER: 60/231,397
;; PRIOR FILING DATE: 2000-09-08

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2005, 22:55:58 ; Search time 133.592 Seconds
(without alignments)
5181.022 Million cell updates/sec

Title: US-09-972-916b-6

Perfect score: 423
Sequence: 1 catggcgccagcggcgacac.....ttccgctactagtagccgc 423

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/6C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.8	14.4	9173	4 US-09-949-001-30	Sequence 30, Appl
2	60.8	14.4	9174	4 US-09-949-001-36	Sequence 36, Appl
3	55.2	13.0	194	4 US-08-945-140-1	Sequence 1, Appl
4	55.2	13.0	13011	2 US-08-791-849A-14	Sequence 14, Appl
5	44.8	10.6	40586	4 US-09-949-016-16965	Sequence 16965, A
6	41	9.7	767677	4 US-09-949-016-12147	Sequence 12147, A
7	41	9.7	767677	4 US-09-949-016-17361	Sequence 17361, A
8	40.2	9.5	15108	4 US-09-949-016-11786	Sequence 11786, A
9	40.2	9.5	15108	4 US-09-949-016-17205	Sequence 17205, A
10	38.4	9.1	11009	4 US-09-845-583A-1	Sequence 1, Appl
11	37.2	8.8	35471	4 US-09-949-016-12957	Sequence 12957, A
12	36.8	8.7	30244	4 US-09-949-016-12208	Sequence 12208, A
13	36.8	8.7	30245	4 US-09-949-016-13550	Sequence 13550, A
14	36.6	8.7	3054	3 US-09-484-970B-138	Sequence 138, App
15	36.4	8.6	134008	4 US-09-949-016-13841	Sequence 13841, A
16	36	8.5	16520	4 US-09-949-016-14710	Sequence 14710, A
17	36	8.5	16520	4 US-09-949-016-15394	Sequence 15394, A
18	35.4	8.4	1503	4 US-09-902-540-1970	Sequence 1970, Ap
19	35.4	8.4	16423	4 US-09-902-540-1120	Sequence 1120, Ap
20	35	8.3	601	4 US-09-949-016-19194	Sequence 19194, A
21	35	8.3	601	4 US-09-949-016-73784	Sequence 73784, A
22	35	8.3	2293	4 US-09-949-016-2153	Sequence 2153, Ap
23	35	8.3	2408	4 US-09-949-016-74	Sequence 74, Appl
24	35	8.3	7561	4 US-09-949-016-11816	Sequence 11816, A
25	35	8.3	7562	4 US-09-949-016-13895	Sequence 13895, A
26	34	8.0	749	3 US-09-257-583-12	Sequence 12, Appl
27	33.8	8.0	601	4 US-09-949-016-86816	Sequence 86816, A

c 28	33.8	8.0	601	4 US-09-949-016-86817	Sequence 86817, A
c 29	33.8	8.0	601	4 US-09-949-016-86818	Sequence 86818, A
c 30	33.8	8.0	601	4 US-09-949-016-86819	Sequence 86820, A
c 31	33.8	8.0	601	4 US-09-949-016-86820	Sequence 86821, A
c 32	33.8	8.0	601	4 US-09-949-016-86821	Sequence 86822, A
c 33	33.8	8.0	601	4 US-09-949-016-86822	Sequence 86823, A
c 34	33.8	8.0	247781	4 US-09-949-016-14193	Sequence 14193, A
c 35	33.6	7.9	601	4 US-09-949-016-21962	Sequence 21962, A
c 36	33.6	7.9	601	4 US-09-949-016-21963	Sequence 21963, A
c 37	33.6	7.9	601	4 US-09-949-016-84283	Sequence 84283, A
c 38	33.6	7.9	601	4 US-09-949-016-84284	Sequence 84284, A
c 39	33.6	7.9	14115	4 US-09-949-016-17490	Sequence 17490, A
c 40	33.6	7.9	455726	4 US-09-949-016-14157	Sequence 14157, A
c 41	33.6	7.9	481115	4 US-09-949-016-11940	Sequence 11940, A
c 42	33.4	7.9	601	4 US-09-949-016-35649	Sequence 35649, A
c 43	33.4	7.9	601	4 US-09-949-016-177807	Sequence 177807, A
c 44	33	7.8	57978	4 US-09-949-016-16667	Sequence 16667, A
c 45	32.8	7.8	1503	4 US-09-902-540-1970	Sequence 1970, Ap

ALIGNMENTS

```

RESULT 1
US-09-949-001-30
; Sequence 30, Application US/09949001
; Patent No. 6825336
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO00789
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 9173
; TYPE: DNA
; ORGANISM: Human
US-09-949-001-30

Query Match      14.4%; Score 60.8; DB 4; Length 9173;
Best Local Similarity 66.5%; Pred. No. 2e+09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

QY 206 CACAGCAAAACAACTTATTGTAACAGGGGATCTAGCACGCTGCCCTGACAAATCAT 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1877 CACTAGCAAAACAACTTATTGTAACACTAGCTCTAGCGCTGCGGCTGCCAATCAT 1936
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 TAACCC-----GTGCTCCGAGCCAGGCGCTTCAATGAGCCCTGGATGCGCAGC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1937 TAACTCTCTGAGCAAGTGGCGGCTGTGCTTTTAAGTGTGCGCTGTCTCAGC 1996
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 CAGCATGTGTCATGTCGCCCGGAGACACAAACGAGCATTTGAACACTGC-ACACGG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1997 GAGCATGCGGCGCACCGCATTC-----CATTCAGCAGCATCTGCGCGCGCGCGG 2048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 CCATCTGCCGAGAGCTGTGACCACTTC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2049 CCACCTCCGAGAGGAGCAGCTGCCACCGCTCC 2080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-09-949-001-36
; Sequence 36, Application US/09949001
; Patent No. 6825336
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH OSTEOPOROSIS, METHODS OF DETECTION AND USES THEREOF

```

```

; FILE REFERENCE: CL000789
; CURRENT APPLICATION NUMBER: US/09/949,001
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/231,323
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 36
; LENGTH: 9174
; TYPE: DNA
; ORGANISM: Human
; US-09-949-001-36

Query Match
Best Local Similarity 66.5%; Pred. No. 2e-09;
Matches 141; Conservative 0; Mismatches 52; Indels 19; Gaps 3;

14.4%; Score 60.8; DB 4; Length 9174;

QY 206 CACAGCAAAACAACTTATTTGAAACAGGGGATCTAGACAGCTGCGCTGACATCAT 265
    |||
DB 1877 CACTAGCAAAACAACTTATTTGAAACACTCAGCTCTAGCGTGCGCGCTGCAATCAT 1936
    |||

QY 266 TAACCC-----GTGCTGCCGAGCAGCCCTTCATTAAGCCCTGGGTATGCGCAGC 315
    |||
DB 1937 TAACCTCTGTGCAAGTGCGCGGCTGTGCTTTATAGGTGCGCGCTGTGTCAGC 1996
    |||

QY 316 CAGCATGTGCACTGCGCGCGGAGACAAACCCAGCAGCATTTGAACACTGC-ACAGG 374
    |||
DB 1997 GAGCATGCGCAGCCGCTATCC-----CATTCAGCAGCAATCTGCCCGCGCGCGC 2048
    |||

QY 375 CCATCTGCCAGAGAGCTGTGACCAACCATCTC 406
    |||
DB 2049 CCACCTCCCAAGAGCACTGGCGCACTGCC 2080
    |||

```

```

RESULT 3
US-08-945-140-1
; Sequence 1, Application US/08945140
; Patent No. 6309878
; GENERAL INFORMATION:
; APPLICANT: CHEN, Ruihuan
; APPLICANT: DOIRON, Bruno
; APPLICANT: KAHN, Axel
; TITLE OF INVENTION: GLUCOSE-INDUCIBLE RECOMBINANT VIRAL
; TITLE OF INVENTION: VECTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,140
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/04558
; FILING DATE: 14-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR96/00560
; FILING DATE: 12-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky Esq., Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: EX95002-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3816

```

```

; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-945-140-1

Query Match
Best Local Similarity 88.2%; Pred. No. 2.4e-08;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

13.0%; Score 55.2; DB 3; Length 194;

QY 1 CATGGGCGCAGGGGCACTCCCGTGTCTTCTGACTCTGCCCCCAGTACATGGGCGC 60
    |||
DB 11 CATGGGCGCAGGGGCACTCCCGTGTCTTCTGACTCTGCCCCCAGTACATGGGCGC 70
    |||

QY 61 ACGGGGCA 68
    |||
DB 71 CGTTGGCA 78
    |||

```

```

RESULT 4
US-08-791-849A-14
; Sequence 14, Application US/08791849A
; Patent No. 5914449
; GENERAL INFORMATION:
; APPLICANT: Makoto MURASE et al.
; TITLE OF INVENTION: Method for Increasing Storage
; TITLE OF INVENTION: Lipid Content in Plant Seed
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Penack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,849A
; FILING DATE: January 30, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELE:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: rat (Rattus norvegicus)
; FEATURES:
; NAME/KEY: CDS
; LOCATION: join(3212..3218, 3766..3948, 5917..6008,
; LOCATION: 6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077

```



```
LOCATION: 9298..9479, 10163..10269)
FEATURE:
NAME/KEY: exon
LOCATION: join(3194..3218, 3766..3948, 5917..6008,
6152..6283, 6418..6604, 6921..7191, 7302..7452, 7665..7817, 7911..8077)
FEATURE:
NAME/KEY: intron
LOCATION: join(3219..3765, 3949..5916, 6009..6151,
6284..6417, 6605..6920, 7191..7301, 7453..7664, 7818..7910, 8078..9297)
LOCATION: 9480..10162)
US-08-791-849A-14
```

```
Query Match 13.0%; Score 55.2; DB 2; Length 13011;
Best Local Similarity 88.2%; Pred. No. 1.6e-07;
Matches 60; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTACATGGGCGC 60
DB 3021 CATGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTACAGGCTTC 3080
QY 61 ACGGGGCA 68
DB 3081 CGTTGGCA 3088
```

```
RESULT 5
US-09-949-016-16965/C
Sequence 16965, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16965
LENGTH: 40586
TYPE: DNA
ORGANISM: Human
US-09-949-016-16965
```

```
Query Match 10.6%; Score 44.8; DB 4; Length 40586;
Best Local Similarity 51.5%; Pred. No. 0.00073;
Matches 103; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
QY 4 GGGCGCAGCGGCGACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTACATGGGCGCAG 63
DB 26135 GGGGTCGCTGGGAGGAGGTGGTGCAGGAGGTGTCCCACTGTCTCAGGGGCGAG 26076
QY 64 GGGCACTCCCGTGTCTCTGACTCTGAGCCGCCAGTGTACATGGGCGCAGGGGCACTCC 123
DB 26075 GGGGAGGAGAGGTGGTGCAGGAGGTGTCCCACTGTCTCAGGGGCTGGGGGAGGGA 26016
QY 124 CGTGTCTCTGACTCTGAGCCGCCAGTGTACATGGGCGCAGGGGCACTCCCGTGTCTC 183
DB 26015 GGTGGTGCAGGTGAGGTGTCCCACTGTCTGAGGGGCTGGGGGAGGAGGAGTGGTGC 25956
QY 184 TGAAGTGTGAGCCGCCAGTGT 203
DB 25955 AGGCGAGGTGTCCCACT 25936
```

```
RESULT 6
US-09-949-016-12147
```

```
Sequence 12147, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12147
LENGTH: 76767
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(76767)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147
```

```
Query Match 9.7%; Score 41; DB 4; Length 76767;
Best Local Similarity 51.4%; Pred. No. 0.05;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
```

```
QY 14 GGCATCCCCGTTCTCTGACTCTGAGCCGCCAGTGTACATGGGCGCAGGGGCACTCCC 73
DB 562328 GGGAGTCCCGTGTCTCTGCTCAGGGCCCGTAGTACTGTAGGGAGTGGAGTCCC 562387
QY 74 GGTGTCTGAGCTCTGAGCCGCCAGTGTACATGGGCGCAGGGGCACTCCCGTGTCTC 133
DB 562388 GTGATCTCAGACTAGAGGCTCCGTGTGACTGTGGGTGGGAGTCCGTGATCTCC 562447
QY 134 GGAAGTCCCGCCAGTGTACATGGGCGCAGGGGCACTCCCGTGTCTGAGCTCTGG 193
DB 562448 TTGCTCAGGGCCCGGTGAGACTGTGGGTGAGGAGTCCATGATCTCTGCTCAGG 562507
QY 194 CCCCC 198
DB 562508 GCCCC 562512
```

```
RESULT 7
US-09-949-016-17361
Sequence 17361, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17361
LENGTH: 76767
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
```

/ LOCATION: (1)...(767677)
 / OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17361

Query Match
 Best Local Similarity 9.7%; Score 41; DB 4; Length 767677;
 51.4%; Pred. No. 0.05;
 Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 14 GGCACCTCCGCTGTTCTTGAGACTGTCGGCCCACTGATACATGAGGCGCACAGGCGCACTGCC 73
 DB 562328 GGGAGTCCCGTATCTCTTCTGCTCAGAGGCCCGTGATGTAAGGTGTAGAGAGTCCC 562387
 QY 74 GTGATCTCGACTGTGCGCCCGAGTGTATGATGAGCGCGCACTCCCGTGTTCCT 133
 DB 562388 GTGATCTCGACTGTGCGCGCTCGTGTGACTGTGCGGAGTCCCGTGTATCTCC 562447
 QY 134 GGAATCTGCGCCCGCACTGTACATGAGCGCGCACAGGCGCACTCCCGTGTTCCTGACTCGG 193
 DB 562448 TTGCTCAGAGGCGCCCGTGAAGACTGTGGTGTAGGAAATCCATGATCTCTTGTCTCAGG 562507
 QY 194 CCCCC 198
 DB 562508 GCCCC 562512

RESULT 8
 US-09-949-016-11786

/ Sequence 11786, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016
 / CURRENT FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 11786
 / LENGTH: 15108
 / TYPE: DNA
 / ORGANISM: Human
 US-09-949-016-11786

Query Match
 Best Local Similarity 9.5%; Score 40.2; DB 4; Length 15108;
 78.7%; Pred. No. 0.015;
 Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 CACGGGGCACTCCCGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCACAGGCGCA 68
 DB 2357 CACGGGGCACTCCCGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCGA 2416
 QY 69 C 69
 DB 2417 C 2417

RESULT 9
 US-09-949-016-17205

/ Sequence 17205, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:
 / APPLICANT: VENTER, J. Craig et al.
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 / FILE REFERENCE: CL001307
 / CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14
 / PRIOR APPLICATION NUMBER: 60/241,755
 / PRIOR FILING DATE: 2000-10-20
 / PRIOR APPLICATION NUMBER: 60/237,768
 / PRIOR FILING DATE: 2000-10-03
 / PRIOR APPLICATION NUMBER: 60/231,498
 / PRIOR FILING DATE: 2000-09-08
 / NUMBER OF SEQ ID NOS: 207012
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 17205
 / LENGTH: 15108
 / TYPE: DNA
 / ORGANISM: Human
 US-09-949-016-17205

Query Match
 Best Local Similarity 9.5%; Score 40.2; DB 4; Length 15108;
 78.7%; Pred. No. 0.015;
 Matches 48; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 9 CACGGGGCACTCCCGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCACAGGCGCA 68
 DB 2357 CACGGGGCACTCCCGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCGA 2416
 QY 69 C 69
 DB 2417 C 2417

RESULT 10
 US-09-845-583A-1/c

/ Sequence 1, Application US/09845583A
 / Patent No. 6635616
 / GENERAL INFORMATION:
 / APPLICANT: Bungeon, Robert
 / APPLICANT: Brunken, William Joseph
 / APPLICANT: Champilaud, Marie-France
 / APPLICANT: Hunter, Dale
 / TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
 / FILE REFERENCE: 10287-056001
 / CURRENT APPLICATION NUMBER: US/09/845,583A
 / CURRENT FILING DATE: 2001-04-30
 / PRIOR APPLICATION NUMBER: US 60/200,863
 / PRIOR FILING DATE: 2000-05-01
 / NUMBER OF SEQ ID NOS: 18
 / SOFTWARE: FastSeq for Windows Version 4.0
 / SEQ ID NO 1
 / LENGTH: 11009
 / TYPE: DNA
 / ORGANISM: Mus musculus
 US-09-845-583A-1

Query Match
 Best Local Similarity 9.1%; Score 38.4; DB 4; Length 11009;
 52.5%; Pred. No. 0.052;
 Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 9 CACGGGGCACTCCCGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCACAGGCGCA 68
 DB 6164 CCGGGGAGACAGGTGCACTGCGCCCGTGTGTGATCAGTGGCTCCGGGAGACATGGGAG 6105
 QY 69 CTCCTGTGTCTCTGACTGTGCGCCCGCACTGTACATGAGCGCGCACAGGCGCACTCCGTGG 128
 DB 6104 CGCTGGAGCCCTTCTGTGGAGGCCCGCACTGAGCGAGCGGCGCACTGAGGCACTGGGGT 6045
 QY 129 TTCTGGAATCTGGCCCCCACTGTACATGAGCGCGCACAGGCGG 168
 DB 6044 CCGTGTGCTCTGTGCGCACTGACATGACGACCGCTTGAG 6005

RESULT 11
 US-09-949-016-12957

/ Sequence 12957, Application US/09949016
 / Patent No. 6812339
 / GENERAL INFORMATION:

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12957
; LENGTH: 35471
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12957

Query Match
Best Local Similarity 8.7%; Score 37.2; DB 4; Length 35471;
Matches 96; Conservative 0; Mismatches 68; Indels 8; Gaps 1;

QY 20 CCGGTGTTCTGACTGAGCCGCGGAGTGTACATGAGGCGGCGACT-----C 71
DB 22548 CCGGTGTTCTGAGTGTGAGGAGGAGTGTGCTGCGGAGGAGTCCGCGTCCCGTGGC 22607

QY 72 CCGGTGTTCTGACTGAGTGTGAGGAGGAGTGTGCTGCGGAGGAGTCCGCGTGGC 131
DB 22608 CCGGTGTTCTGAGTGTGAGGAGGAGTGTGCTGCGGAGGAGTCCGCGTGGC 22667

QY 132 CTGGACTGTGCGCCGAGTGTACATGAGGCGGCGACTCCCGTGTTC 183
DB 22668 CTGGACTGTGCTCCCTGAGGAGTGTGCTGCGGAGGAGTCCCGTGTTC 22719

RESULT 12
US-09-949-016-12208
; Sequence 12208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12208
; LENGTH: 30244
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12208

Query Match
Best Local Similarity 8.7%; Score 36.8; DB 4; Length 30244;
Matches 92; Conservative 0; Mismatches 72; Indels 2; Gaps 1;

QY 40 GCCCCAGTGTACATGAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGT 99
DB 13568 GCCAGCTGTACATCTGGTGTCTGAGGAGCCCGGCGGTTTCACGCGATTTGCCCTGGGT 13627
```

```

QY 100 GTACATGAGGCGGAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGTATGAG 159
DB 13628 ATTCAGGAGACCGGCGGAGTGTGAGGAGTCT--ACAGCAGCGAGTGTACATCTGG 13685

QY 160 CGCAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGTATGAG 205
DB 13686 TGCTGTGGGAGCCCGGAGTGTTCATGAGTGTGCGGTTT 13731

RESULT 13
US-09-949-016-13550
; Sequence 13550, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13550
; LENGTH: 30245
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13550

Query Match
Best Local Similarity 8.7%; Score 36.8; DB 4; Length 30245;
Matches 92; Conservative 0; Mismatches 72; Indels 2; Gaps 1;

QY 40 GCCCCAGTGTACATGAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGT 99
DB 13568 GCCAGCTGTACATCTGGTGTCTGAGGAGCCCGGCGGTTTCACGCGATTTGCCCTGGGT 13627

QY 100 GTACATGAGGCGGAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGTATGAG 159
DB 13628 ATTCAGGAGACCGGCGGAGTGTGAGGAGTCT--ACAGCAGCGAGTGTACATCTGG 13685

QY 160 CGCAGGCGGAGTCCCGTGTCTGAGTGTGCGCCCGCAGTATGAG 205
DB 13686 TGCTGTGGGAGCCCGGAGTGTTCATGAGTGTGCGGTTT 13731

RESULT 14
US-09-949-970B-138
; Sequence 138, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkman, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO: 138
; LENGTH: 3054
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 336987.1CB1
```

US-09-484-970B-138

Query Match	8.7%	Score 36.6;	DB 3;	Length 3054;
Best Local Similarity	62.6%;	Pred. No. 0.11;		
Matches 57;	Conservative 0;	Mismatches 34;	Indels 0;	Gaps 0;

QY 11 CGGGGACCTCCCGTGGTTCCTGGACTCTGCGCCCAAGTGTACATGGGCGCACCGGGGACT 70
Db 265 CCGGGCTCTCCCTTGGCTCCAGCCTTAGGCCACCGGCTTGGATGAGCGCTCCGAGGCTA 324

Oy 71 CCCGCGTTCCTGACTCTGGCCCCCAGTGT 101
||| ||| ||| ||| |||
Db 325 CCCGTCCTCCAGGGGTCTTGCTTCCGAAGT 355

RESULT 15
US-09-949

US-09-949-016-13841
; Sequence 13841, Application US/09949016

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: IIS/09/949 016

```

;; CURRENT FILING DATE: 2000-04-14

; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03 ;

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

```

; NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0

```

; SOFTWARE: FAST
; SEO ID NO 13841

LENGTH: 134008

TYPE: DNA

ORGANISM: Human

FEATURE:

```

; NAME/KEY: misc_feature
; LOCATION: (1) (1234000)

```

LOCATION: (1) :..(134008)

OTHER INFORMATION
IS-09-949-016-13841

Query Match 8.6%; Score 36.4; DB 4; Length 134008;

Matches	85;	Conservative	0;	Mismatches	81;	Indels	0;	Gaps	0;
35	CTCTGGCCCCAGTGTACATGGCGCCACGGGGCACTCCCGTGTCTCTGGACTCTGGCCC	94							

Db	127072	CACGGACACCAGCGCCGACGCACAGCAGCTGTGACGCC	127
QY	95	CCAGTGTACATGGGCGCACGGGGCACTCCCGTTTCCTGACTTTGGCCCCCAGTGAC	154

Db 127132 ACAGATGACGTGACCGCACGACGATGATGCATGGACGACGCTGGACCGCAC

Search completed: September 1, 2005, 07:07:22
Job time : 139.759 secs